

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 03:30:35 ; Search time 1713.89 Seconds  
(without alignments)  
3709.145 Million cell updates/sec

Title: US-09-525-361A-23

Perfect score: 471

Sequence: 1 ctttgaagcatttttctgtg.....aactatgagcgagtaacat 471

Scoring table:

IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_hic:\*
- 9: gb\_estl:\*
- 10: gb\_est2:\*
- 11: gb\_hic:\*
- 12: gb\_gss:\*
- 13: em\_gss\_hum:\*
- 14: em\_gss\_inv:\*
- 15: em\_gss\_pln:\*
- 16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	431.8	91.7	436	10	W72837
2	403.2	85.6	432	10	W72838
3	330.8	83.0	420	9	AW070916
4	388.6	82.5	423	10	BF742809
5	379	80.5	402	9	AI139456
6	345.2	73.3	464	10	BG202312
7	331	70.3	376	10	BG218084
8	316.8	67.3	345	9	BE092421
9	309.8	65.8	391	10	BE185269
10	308.8	65.6	403	10	R75793
11	301.4	64.0	314	12	AZ694036
12	299.2	63.5	343	10	BG202313
13	294	62.4	313	9	AW176044
14	289.4	61.4	294	9	AA340069
15	281	59.7	316	9	BE186013
16	281	59.7	316	10	BE926938
17	255.8	54.3	443	9	AI127172

18	247.6	52.6	308	9	AI905687
19	217.8	46.2	251	10	BE197186
20	202.8	43.1	221	9	BE074512
21	197.8	42.0	462	9	AW294149
22	197	41.8	224	9	AW291950
23	196.6	41.7	230	10	BE815819
24	187.8	39.9	248	10	BE815824
25	159	33.8	193	9	AI905624
26	154	32.7	224	9	AI905837
27	151	32.1	211	10	BE183176
28	150	31.8	212	10	BE192597
29	149.4	31.7	178	9	AI905623
30	140.4	29.8	211	10	BG207535
31	133	28.2	136	9	AI905633
32	102.8	21.8	224	10	BE815853
33	85	18.0	498	10	BE1279448
34	82.6	17.6	536	10	BI339750
35	82.6	17.5	572	10	BF080542
36	82.4	17.5	490	10	BI279279
37	82.4	17.5	492	10	BI279431
38	82.4	17.5	499	10	BI279473
39	81.8	17.4	486	10	BI282803
40	80.8	17.2	481	10	BI279272
41	80.8	17.2	490	10	BI279159
42	80.8	17.2	493	10	BI279513
43	80.8	17.2	493	10	BI279518
44	80.8	17.2	495	10	BI279214
45	80.8	17.2	496	10	BI279121

#### ALIGNMENTS

RESULT 1

W72837

LOCUS

DEFINITION

zds57g12.r1 Soares\_fetal\_heart\_NbHH19W Homo sapiens cDNA clone

IMAGE:344806 5' similar to contains element MER40 repetitive

element ;, mRNA sequence.

W72837

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1

436 bp

linear

EST 16-OCT-1996

W72837

zds57g12.r1 Soares\_fetal\_heart\_NbHH19W Homo sapiens cDNA clone

IMAGE:344806 5' similar to contains element MER40 repetitive

element ;, mRNA sequence.

W72837

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1

436 bp

linear

EST 16-OCT-1996

W72837

zds57g12.r1 Soares\_fetal\_heart\_NbHH19W Homo sapiens cDNA clone

IMAGE:344806 5' similar to contains element MER40 repetitive

element ;, mRNA sequence.

W72837

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1

436 bp

linear

EST 16-OCT-1996

W72837

zds57g12.r1 Soares\_fetal\_heart\_NbHH19W Homo sapiens cDNA clone

IMAGE:344806 5' similar to contains element MER40 repetitive

element ;, mRNA sequence.

W72837

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1

436 bp

linear

EST 16-OCT-1996

W72837

zds57g12.r1 Soares\_fetal\_heart\_NbHH19W Homo sapiens cDNA clone

IMAGE:344806 5' similar to contains element MER40 repetitive

element ;, mRNA sequence.

W72837

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1

436 bp

linear

EST 16-OCT-1996

W72837

zds57g12.r1 Soares\_fetal\_heart\_NbHH19W Homo sapiens cDNA clone

IMAGE:344806 5' similar to contains element MER40 repetitive

element ;, mRNA sequence.

W72837

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1

436 bp

linear

EST 16-OCT-1996

W72837

zds57g12.r1 Soares\_fetal\_heart\_NbHH19W Homo sapiens cDNA clone

IMAGE:344806 5' similar to contains element MER40 repetitive

element ;, mRNA sequence.

W72837

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1

436 bp

linear

EST 16-OCT-1996

W72837

zds57g12.r1 Soares\_fetal\_heart\_NbHH19W Homo sapiens cDNA clone

IMAGE:344806 5' similar to contains element MER40 repetitive

element ;, mRNA sequence.

W72837

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1

436 bp

linear

EST 16-OCT-1996

W72837

zds57g12.r1 Soares\_fetal\_heart\_NbHH19W Homo sapiens cDNA clone

IMAGE:344806 5' similar to contains element MER40 repetitive

element ;, mRNA sequence.

W72837

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1

436 bp

linear

EST 16-OCT-1996

W72837

zds57g12.r1 Soares\_fetal\_heart\_NbHH19W Homo sapiens cDNA clone

IMAGE:344806 5' similar to contains element MER40 repetitive

element ;, mRNA sequence.

W72837

VERSION

KEYWORDS

SOURCE

ORGANISM



```
RESULT 3
AW070916/c 420 bp mRNA linear EST 20-OCT-2000
LOCUS xaj1001.x1 NCI_CGAP_Brl8 Homo sapiens cDNA clone IMAGE:2568385 3'
DEFINITION similar to contains element TARI repetitive element ;, mRNA
sequence.
ACCESSION AW070916
VERSION AW070916.1 GI:6025914
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 420)
JOURNAL NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps@mail.nih.gov
Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life
Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The
I.M.A.G.E. Consortium DNA Sequencing by: Washington University
Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/hbrp/image/image.html
Insert Length: 545 Std Error: 0.00
Seq primer: -40UP from Gbco.
FEATURES
Location/Qualifiers
1..420
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2568385"
/clone_lib="NCI_CGAP_Brl8"
/tissue_type="four pooled high-grade tumors, including two
primary tumors and two metastatic to ovary"
/lab_host="PH10B"
/note="Organ: breast; Vector: pcMV-SPORT6; Site_1: Salt;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies."
BASE COUNT 119 a 76 c 116 g 109 t
ORIGIN
source
Query Match 83.0%; Score 390.8; DB 9; Length 420;
Best Local Similarity 98.1%; Pred. No. 2.8e-90;
Matches 406; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
QY 57 tagcagctcctggtactcttggagtttccatcttctgtctctccacgaatccgacaa 116
|||||
Db 420 TACGACTCCTGGTACTCTTGGGAGTTTCCATCTTCTGCTCTGCCCGAATCCGACAA 361
|||||
QY 117 cagctgctccagctgacagctaccagctactggctcgtgctgctgatgatgaagccctgatg 176
|||||
Db 360 CAGCTGCTCCAGCTGACAGCTATCCAGCTACTGGTCTCTGCTGATGATGAAGCCCTGATG 301
|||||
QY 177 ctgaaacacactgctgtgcaaacactggaacactgctgctctaccactgcaacacag 236
|||||
Db 300 CTGAACACCACTGCTGTGCAACCACTGCCAGCACTGCTGCTCTCTACCACTGCAACACCG 241.
|||||
QY 237 ctgctctaccactgctgtgaaagacatccagttttacccaaagggttggggtctcc 296
|||||
Db 240 CTGCTTCTACCACTGCTCTGAAGACATTCAGTTTACCCAAATGGGTGGGATCTCC 181
|||||
QY 297 cgaatggtagtggtgtccctgagatggaatcagcttgagttctctgcaattgggtcaca 356
|||||
Db 180 GCAATGGTAGAGTGTCTCCCTGAGATGGAATCAGCTTGAGTTCTTGCAAATT-GGTACA 122
|||||
QY 357 actattcagctctcgtgatttccaaactactactacccttgcctcagcatcccccttta 416
|||||
```

```
Db 121 ACTATTTCATGCTTCCTGTTGATTTCATCCAACTACTTACCTTCCTTCAGATATCCCTTTTA 62
QY 417 tctctaatacagtttatttcttcttcaataaaaaaactatgagcgcagtaaca 470
|||||
Db 61 TCTCTAATCAGTTTATTTCTTCTTCAATAAAAAATAACTATGAGCAACAAAAA 8
|||||
RESULT 4
BF742809 423 bp mRNA linear EST 10-JAN-2001
LOCUS IL2-BT0807-041000-176-G03 BT0807 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF742809
VERSION BF742809.1 GI:12069485
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 423)
JOURNAL Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
MEDLINE Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
20202663 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
COMMENT Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shogun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=IL2&l2=IL2-BT0807-
041000-176-G03&l3=2000-10-04&l4=1)
Seq primer: puc 18 forward
High quality sequence stop: 421.
FEATURES
Location/Qualifiers
1..423
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BT0807"
/dev_stage="Adult"
/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 89 a 123 c 87 g 124 t
ORIGIN
source
Query Match 82.5%; Score 388.6; DB 10; Length 423;
Best Local Similarity 98.8%; Pred. No. 1e-89;
Matches 402; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 1 ctttgaagcatttttctgtgctccctgatcttcaggtcacccaccatgaagttcttagc 60
|||||
Db 15 CTTTGAAGCATTTTGTCTGCTGCTCCCTGATCTTCAGTCCACCACCATGAAGTCTTAGC 74
|||||
QY 61 agtctcgttactcttggagtttccatcttctggtctctgccagaatccgcacaacgc 120
|||||
Db 75 AGTCTGTGCTACTCTTGGGAGTTTCCATCTTTCTGTGCTCTGCCAGAAATCCGACAACAGC 134
|||||
QY 121 tgcctcagctgacagctatccagctactggtcctcgtgatgatgaagccctcgtgatg 180
|||||
```

```
Db 135 TGCTCCAGCTGACACGATATCCAGCTACTGGTCTGCTGATGATGAAGCCCTCGATGTTGA 194
Oy 181 aaccactgctgctgcaacaactgcgaccactgctgctctaccactgcaaccacgcgtgc 240
Db 195 AACCACTGCTGCTGCAACCACTGCGACCACTGCTGCTCTACCACTGCAACCACTGCTGC 254
Oy 241 tctaccactgctcgtlaagagacattccagttttaccacaaatgggtgggagatctcccgaa 300
Db 255 TTCTACCACTGCTGCTGAAGACATTCACGTTTACCCAAATGGGTGGGGATCTCCCGAA 314
Oy 301 tggtagagtgctccctgagatggaatcagcttgagctctctgcaattgggtgcacaacta 360
Db 315 TGCTAGAGTCTGCTCCCTGAGATGGAATCAGCTTGAGCTCTTCTGCAATT-GGTACAACTA 373
Oy 361 ttcactctctctgctgatttcacacactacttaccttgcctacgata 407
Db 374 TTCTACGCTCTCTGATTTTCATCCAACTACTTACCTTGCTCTACACTA 420

RESULT 5
A1139456/c
LOCUS
DEFINITION
q20g01.x1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone
IMAGE:1710192 3' similar to contains element MER40 repetitive
element ;, mRNA sequence.
ACCESSION
A1139456
VERSION
A1139456.1 GI:3645428
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 402)
AUTHORS
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL
Contact: Robert Strausberg, Ph.D.
COMMENT
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LNL : contact the
IMAGE Consortium (info@image.lni.gov) for further information.
Insert Length: 510 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 398.
Location/Qualifiers
1..402
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1710192"
/clone_lib="Soares_fetal_heart_NbHH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: heart; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer
TGTTACCAATCTGAAGTGGAGCGCGCATCTTTTCTTTTCTTTT 3'),
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot - 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NbHH19W."
BASE COUNT 115 a 72 c 115 g 100 t
ORIGIN

Query Match
Best Local Similarity 80.5%; Score 379; DB 9; Length 402;
Matches 393; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
```

```
Oy 65 ctggtactcttggggagtttccatcttcttggctctgcccagaatccgacaacagctgct 124
Db 402 CTGGTACTCTTGGGAGTTTCCATCTATCTGGTCTCTGCCAGAAATCCGACAACAGCTGCT 343
Oy 125 ccagctgacacgactccagctactggtctgctgctgatgagccctgatgctgaaacc 184
Db 342 CCAGCTGACACGATTCACGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 283
Oy 185 actgctgctgcaacaactgcgaccactgctgctctaccactgcaaccacgcgtcttct 244
Db 282 ACTGCTGCTGCAACCACTGCGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 223
Oy 245 accactgctgtaagacattccagttttaccacaaatgggtgggagatctcccgaaatgg 304
Db 222 ACCACTGCTGCTGTAATTCATCCCAACTACTTACCTTGCTGCTAGGATATCCCTTTATCTCTAAT 44
Oy 425 cagttattttctttccaaataaaaaataactatgagcga 463
Db 43 CAGTTTATTTTCTTTCAATAAAAAATAACTATGAGCAA 5

RESULT 6
BG202312
LOCUS
DEFINITION
RST21668 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION
BG202312
VERSION
BG202312.1 GI:13723999
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 464)
AUTHORS
Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,
Lerner,B., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith
,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
,J., Danzig,J. and Ducar,M.
Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)
21227151
Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com
High quality sequence stop: 464.
Location/Qualifiers
1..464
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="Hr1080"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is Hr1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in Hr1080 under normal circumstances."
BASE COUNT 126 a 119 c 91 g 128 t
ORIGIN
```

BASE COUNT	102 a	103 c	67 g	104 t	expressed in HT1080 under normal circumstances
ORIGIN					
Query Match					
Best Local Similarity					
Matches 351:	Conservative	0;	Mismatches	15;	Indels 1;
Qy	104	cagaatccgacacagctgcgccagctgcacagctalccagctactggtcctgctgtagta	Score 331;	DB 10;	Length 376;
Db	4	CAGAGATCCACACAGCTGCTCCAGTGACAGCTATCCAGCTACTCGTCTCTCTATCA	70.3%;		
			95.6%;		
Qy	164	gaagccctgatcctgaaccactgctgctgcacacactggacacactgctgctcctactac	Pred. No. 70-75;		
Db	64	GAGCCCTGCTGATCTGAACCCACTGCTGCTGCAACCCTCGCACCACTGCGACCACTGCT			
Qy	224	actgcaaccacgctgcttctaccactgctgctgaagacattccaggtttaccaccaaatg			
Db	124	ACTGCAACCAACCGCTGCTTACCACCTGCTCTGAAGACATTTCCAGTGTTCACCAATG			
Qy	284	gttggggatctccgaatggtagagtggtccctgagatggaatcagctgtgagctctctc			
Db	184	GTTGGGGATCTCCCGAATGGTAGGTGTCTCCCTCAGATGGAATCAGCTTCAGTCTCT			
Qy	344	caattgggtcacaaactatcatgcttctctgtgatttcacaaactacttaocttgcctata			
Db	244	CAATT -GGTCACAACATTTATGCTTCTCTGTGATTTCATCCAACTACTTACCTTGGCCTAT			
Qy	404	gataccctttatctctaactcagtttatctttcttcaataaaaaataactatgagcgt			
Db	303	GATATCCCTTTATCTCTAATCAGTTTATTTTCTTTTCAGATAAAAAATAACTATGAGCA			
Qy	464	gctaaca 470			
Db	363	CAAAAAA 369			

[illegible]

TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE	20202663
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL ( <a href="http://www.ludwig.org.br/scripts/gethtml2.pl?i=4t2-CM1-BT0738-250">http://www.ludwig.org.br/scripts/gethtml2.pl?i=4t2-CM1-BT0738-250</a> ) 400-200-cl24t3-2000-04-254t4+1 Seq primer: puc18 forward

High quality sequence start: 9  
High quality sequence stop: 345.

**FEATURES**

**SOURCE**

```

1. 345
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone_lib="BT0738"
    /dev_stages="Adult"
    /note="Organ: breast; Vector:
    from ORESTES PCR (U.S.
    716 - Ludwig Institute
    into the pUC 18 vector.
    mRNA and cDNA amplification
    stringency conditions."
    location/Qualifiers

```

BASE COUNT	92 a	77 c	103 g	73 t
ORIGIN	emergency conditions.			

Query Match 67.3%; Score 316.8; DB 9; Length 345;  
Best Local Similarity 97.6%; Pred. No. 3.2e-71;  
Matches 332: Conservative 0; Mismatches 7; Indels 1

Qy	37	ggcacacacataaagtctcttaacagctccgtgtaactcttggaagtcttcacatctctcttggt	96
Db	345	ggctaccaccataaagttctttacagctcctggtactctttggagattttccattcttctggt	286
Qy	97	ctctgcccagaatccgacaaacagctgctccagctgacacglatccagctactggtcctgc	156
Db	285	ctctgcccagaatccgacaaacagctgctccacgctcacacgattccacgactactggtcctgc	226
Qy	157	taatgataagccccctgaagtctgaacacactgctctcaacacactgcgacacactgctgc	216
Db	225	tgatgatgataagccccctgatgctgaacacacactgctgctgcaacacactgcgacacactgctgc	166
Qy	217	tcttaccactgcgaacacacccgctgcttaccactgctcgtaaagacattccagttttacc	276
Db	165	tccctaccactgcgaacacacccgctgcttaccactgctgcttaacacactgctcgttaagacattccagttttacc	106
Qy	277	caaatgggttggggatctcccgatgtagtgatgltgctccctgagatggaatcagcttgag	336
Db	105	caaatgggttggggatctcccgatgtagtgatgltgctccctgagatggaatcagcttgag	47
Qy	337	tctctgcaattgggtgcacaactattcatgcttccctgtga	376
Db	46	tcttctgctgctagttgggtcacaaactatttcattgcttccgctgcga	7

RESULT	9	
LOCUS	BG185269	
DEFINITION	R874208 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.	
ACCESSION	BG185269	
VERSION	BG185269.1	GI:13706956
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo;	
	1 (bases 1 to 391)	
REFERENCE	Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,	
AUTHORS	Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,	
	Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith	
	,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher	
	J., Danzig,J. and Ducar,M.	
TITLE	Creation of genome-wide protein expression libraries using random	
	activation of gene expression	
JOURNAL	Nat. Biotechnol. 19 (5), 440-445 (2001)	
MEDLINE	21227151	
COMMENT	Contact: Scott J. Cain	
	Athersys, Inc.	
	3201 Carnegie Ave, Cleveland, OH 44115, USA	
	Tel: 216 431 9900	

**Fax: 216 361 9596**  
**Email: [scain@athersys.com](mailto:scain@athersys.com)**  
**High quality sequence stop: 391.**

[illegible]

```
Location/Qualifiers
1. .391
/organism:"Homo sapiens"
/db_xref:"taxon:9606"
/clone_lib:"Athensys RAGE Library"
/cell_line:"HT1080"
/note:"See 'Creation of Genome-wide Libraries using Random Activation' Nature Biotechnology, in press. No cell type indicated is HT1080, since the method was used, these sequence tags expressed in HT1080 under normal
```

BASE COUNT	105 a	103 c	72 g	101 t	10 others
ORIGIN	expressed in	expressed in	expressed in	expressed in	expressed in
	normal	normal	normal	normal	normal
	circumstances	circumstances	circumstances	circumstances	circumstances

Query Match 65.8%; Score 309.8; DB 10; Length 391;  
Best Local Similarity 97.1%; Pred. No. 2.1e-69;  
Matches 336; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

125	QY	ccagctgacacgtatccagctactggtcctctgctgatagaagagccctcctgatgctgaacc	184
41	Db	CCAGCTGNAACGTATCCAGCTACTGCTCTCTGCTGATGATGAAGCCCTGA-GCTGAAC	99
185	QY	actgctgctgcaacaactgcgaccactgctgctcctaccactgcaaccaccgctgctct	244
100	Db	ACTGCTGCTGCAACCACTGGACCACCTGCTGCTCTTACCACCTGCAACCAACCGTGTCT	159
245	QY	accactgctgtaagaacatccagttcttaccacaaatgggttgggaatctcccgaaaggt	304
160	Db	ACCACCTGCTGTAAGACACATTCAGATTTTACCCAAATGGGTGGGGATCTCCCGAATG	219
305	QY	agagtgctgctcgagatggaaacagcttgagctctctctgcaattgggtggcacaaacttca	364
220	Db	AGAGTGTGTCCTGACATGGAATCAGCTTGAGTCTCTGCAATT-GGTCAACAATATCA	278
365	QY	tgcttctctgattctcatccaactactacccttgctctagatataccccctttatctcta	424
279	Db	TGCTTCTGTGATTTTCATCAACTACTTACCTTTGCCCTAGCATATCCCTTTATCTCTA	338
425	QY	cagtttatttcttccaaataaaaaaactatgagcgagctaaca	470
339	Db	CAGTTATTTCTTCCAAATAAAAAATAACTATGAGCAACAACAAAAA	384

```

RESULT 10
R75793
LOCUS
DEFINITION
IMAGE:138921 5' similar to contains LTR3 repetitive element ; , mRNA
sequence.
ACCESSION R75793
VERSION R75793.1 GI:850475
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 403)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston
,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

```

Email: est@watson.wustl.edu  
Insert Size: 516  
High quality sequence stops: 307  
Source: IMAGE Consortium, LLNL  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 516 Std Error: 0.00  
Seq primer: M13Kp1  
High quality sequence stop: 307.

FEATURES  
source  
1. 403  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="GDB:572980"  
/db\_xref="taxon:9606"  
/clone="IMAGE:158921"  
/clone\_lib="Soares breast 2NBHbst"  
/sex="Female"  
/dev\_stage="adult"  
/lab\_host="DH10B (ampicillin resistant)"  
/note="Organ: breast; Vector: pT7T3D (Pharmacia) with a  
modified polylinker; Site\_1: Not I; Site\_2: Eco RI; lstr  
strand cDNA was primed with a Not I - oligo(dT) primer [5'  
TGTTACCAATCTGAAGTGGAGCGGCCCTTTTTTTTTTTTTTTT 3'],  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of a modified pT7m3 vector (Pharmacia).  
Library went through one round of normalization to a Col -  
230. Library constructed by Bento Soares and M.Fatima  
Bonaldo."

BASE COUNT 82 a 113 c 89 g 116 t 3 others  
ORIGIN  
Query Match 65.6%; Score 308.8; DB 10; Length 403;  
Best Local Similarity 94.0%; Pred. No. 3.7e-69;  
Matches 375; Conservative 0; Mismatches 15; Indels 9; Gaps 5;  
QY 8 gcaattttgtctgtccctgattcttcaggtcaccacattgaagcttcttagcagtcctg 67  
|||||  
Db 2 GCATTTTGTCTGTCTCCCTGATCTTCAGTCCACCACCATGAACTTCTTAGCAGTCCTG 61  
QY 68 gtactcttggaggtttccatcttcttggtctctgcccagaatccgacaacagctgtctca 127  
|||||  
Db 62 GNACTCTTGGGAGTTTCCATCTTCTTGGTCTCTGCCAGAAATCCGACACAGCTGTCTCA 121  
QY 128 qctgacacgtatccagctactgctctgctgatgatgaagccctctgatgctgaacacat 187  
|||||  
Db 122 GCTGACACGTATCCAGTACTGTGCTCTGCTGATGATGAAGCCCTGATGCTGAACCACT 181  
QY 188 gctgctgcaacaactgcgaccactgctgctcctaccactgcaaccacgctgcttacc 247  
|||||  
Db 182 GCTGCTGCAACCACTGGACCACTGCTGCTCTTACCACCTGCAACCACTGCTGCTTACC 241  
QY 248 actgctgtaaaagacttccagttttaccacaaatgggttggggtatcccgatgg--ta 305  
|||||  
Db 242 ACTGCTCGTAAGACATATCCAGTTTACCCAAATGGGTGGGATCTCCGCAATGGGTAG '301  
QY 306 gagtgtctccttag--atggaaatcagtttgagtctt--ctgcaattgggtcacaaacta- 360  
|||||  
Db 302 GACTGTGCTCCCTGAGGATGGGAATCATCAGTTGAGTCTTCTTGGCAATTTGGGTGACAAATAT 361  
QY 361 ttcatgcttctctgtg--atttccatcaactacttaacctt 397  
|||||  
Db 362 TTTCATGCTTCTCTGGATTTCATCCAAATTAATTNACCT 400

RESULT 11  
AZ694036 314 bp DNA linear GSS 18-DEC-2000  
LOCUS  
DEFINITION  
AST-2H8BG1003 Genetrap HL-60 Human Promyelocytic Leukemia Library.  
Homo sapiens genomic 5', DNA sequence.  
ACCESSION  
AZ694036  
VERSION  
AZ694036.1 GI:11878955

KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
GSS.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 314)  
Henkel G., Liyanage, M., Pratt, E., Huang, D., Riley, M., Bernardino, A.,  
Durick, K. and Pollok, B.  
Exon-trap tags from a HL-60 GenomeScreen(TM) Library  
Unpublished (2000)  
Contact: Greg Henkel  
Gene Expression  
Aurora Biosciences Corp.  
11010 Torreyana Road, San Diego, CA 92121, USA  
Tel: 8584048436  
Fax: 8584046719  
Email: henkelg@aurorabio.com  
Pools of cells were isolated from a GenomeScreen(TM) library. The  
library of cells was generated by retroviral integration of a gene  
tagging element consisting of: 1) A promoterless beta-lactamase  
preceded by a splice acceptor as a reporter for gene expression;  
2) A promoter driving neomycin resistance followed by a splice  
donor to trap downstream exons. 3' RACE from neomycin gene was  
performed using total RNA from isolated pools. Output was shotgun  
cloned in pAmp-1 and used to transform DH5-alpha competent  
bacteria. 5' ends of reported sequences were immediately preceded  
by splice donor from the trapping construct.  
Class: exon-trapped.  
Location/Qualifiers  
1. 314  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="Genetrap HL-60 Human Promyelocytic Leukemia  
Library"  
/tissue\_type="acute promyelocytic leukemia"  
/cell\_type="promyeloblast"  
/cell\_line="HL-60"  
/note="Organ: peripheral blood; Vector: pAmp-1; 3' RACE of  
total RNA from genetrap pools; shotgun clone in pAmp-1 and  
used to transform DH5-alpha competent bacteria."  
BASE COUNT 76 a 88 c 56 g 94 t  
ORIGIN

Query Match 64.0%; Score 301.4; DB 12; Length 314;  
Best Local Similarity 99.4%; Pred. No. 2.9e-67;  
Matches 313; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 147 ctggctctgctgatgaagccctctgctgaaccactgctgctgcaacaactggga 206  
|||||  
Db 1 CTGGTCTCTGCTGATGATGAAGCCCTGATGCTGAACCACTGCTGCTGAACCACTGGGA 60  
QY 207 ccactgctgctctaccactgcaaccacccgctgcttctaccactgctgtaaaagacattc 266  
|||||  
Db 61 CCACCTGCTCTCTACCACCTGCAACCAACCGCTGCTTCTACCACCTGCTGTAAGACATTC 120  
QY 267 cagttttaccacaaatgggttggggatctctcccaaaatggtagagtgctgctcctgagatgaa 326  
|||||  
Db 121 CAGTTTATCCCAAAATGGGTGGGGATCTCCCAATGGTAGAGTGTGCTCCTGAGATGAA 180  
QY 327 tcagcttgagctctctgcaattgggtcacaactattcattcctgctgctgattctccaa 386  
|||||  
Db 181 TCAGCTTGAGTCTTCTGCNAATT-GGTCACACTATTCATGCTTCTCTGATTTTCATCAA 239  
QY 387 ctacttaccttgctcacgatatcccccttttatctctaatcagtttattttctttcacaataa 446  
|||||  
Db 240 CTACTTACCTTGCTTACGATATCCCCCTTTATCTCTTAATCAGTTTATTTTCTTCAATAA 299  
QY 447 aaaaataactatgagc 461  
|||||  
Db 300 AAAATAACTATGAGC 314





[illegible]

QY 58 agcagtcctggtactcttgagagatttccatcttcttggtctc-tgccagagaatccgacaa 116  
|||||  
Ub 5 AGCAGTCCTGGTACTCTTGGGAGTTTCCATCTTTCCGGTCTCGTGATCAGAATCCGACAA 64  
|||||  
QY 117 cagctgctccagctgacacgatalccagctactggtcctgctgatgatgaagccccctgag 176  
|||||  
Db 65 CAGCTGCTCCAGCTGACACGTATCCAGCTACGGGTCTCTGCTGATGATGAAGCCCCCTGATG 124  
|||||  
QY 177 ctgaaccactgctgctgcaacaactgcgaccactgctgctcctaccactgcaaccaccg 236  
|||||  
Ub 125 CTGAACACACTGTGCTGCAACCACTGCGACCACCTGCTCTCTACCACTGCAACCAACCG 184  
|||||  
QY 237 ctgcttctaccactgctcglaaagacattccagtttaccacaaatgggttggggatctcc 296  
|||||  
Ub 185 CTGCTTCTACCACTGCTGCTGTAAGACATTCCAGCTTTTACCCAAATGGGTGGGGATCTCC 244  
|||||  
QY 297 cgaatgglagagtgtgtccctgagatgggaatcagcttgagcttctctgcaattgggtcaca 356  
|||||  
Ub 245 CGAATGCTAGAGTGTGTCCCTGAGATGGAATCAGCTTGAGTCTTCTGCAATT-GGTCACA 303  
|||||  
QY 357 actattcatgctt 369  
|||||  
Db 304 ACTATTGATGCTT 316

Search completed: October 5, 2002, 05:07:20  
Job time: 5805 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 03:59:20 ; Search time 46.3 Seconds  
(without alignments)  
2498.777 Million cell updates/sec

Title: US-09-525-361A-23

Perfect score: 471

Sequence: 1 ctttgaagcatttttctgtg.....aactatgagcgagtaacat 471

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA:\*

- 1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	448.4	95.2	497	4	US-09-247-155-41
2	68.6	14.6	397	3	US-09-253-691-3
3	65.4	13.9	477	4	US-09-135-994-1
4	65	13.8	234	1	US-08-469-802B-3
5	65	13.8	234	2	US-08-267-803B-3
6	64.8	13.8	3376	1	US-08-320-559-29
7	64.8	13.8	3376	3	US-08-545-860D-29
8	64.8	13.8	3376	5	PCT-US94-04496-29
9	64.6	13.7	203	4	US-09-043-303-7
10	63.4	13.5	168	1	US-08-469-802B-4
11	63.4	13.5	168	2	US-08-267-803B-4
12	63.4	13.5	171	1	US-08-469-802B-5
13	63.4	13.5	171	2	US-08-267-803B-5
14	63.4	13.5	195	1	US-08-469-802B-2
15	63.4	13.5	195	2	US-08-267-803B-2
16	62.8	13.3	154	1	US-08-469-802B-6
17	62.8	13.3	154	2	US-08-267-803B-6
18	62.2	13.2	165	4	US-09-043-303-17
19	52.6	11.2	2580	3	US-09-050-863-2
20	52.6	11.2	2580	4	US-09-359-081-2
21	52.6	11.2	3489	2	US-08-728-323A-1
22	52.6	11.2	5452	2	US-09-130-114-1
23	52.6	11.2	9600	4	US-08-910-647-1
24	52.6	11.2	10596	1	US-07-884-811-15
25	52.6	11.2	10596	1	US-07-885-971-15
26	52.6	11.2	10596	1	US-08-087-783A-15
27	52.6	11.2	10596	1	US-08-194-088B-15

c 28	52.6	11.2	10596	2	US-08-194-087-15	Sequence 15, Appl
c 29	52.6	11.2	10596	5	PCT-US93-04648-15	Sequence 15, Appl
c 30	52.6	11.2	32207	2	US-08-770-379-20	Sequence 20, Appl
c 31	52.6	11.2	32207	4	US-08-757-669A-20	Sequence 20, Appl
c 32	52.6	11.2	32207	4	US-09-230-371A-20	Sequence 20, Appl
c 33	52.6	11.0	2790	2	US-08-718-661-1	Sequence 1, Appl
c 34	51.2	10.9	1518	2	US-08-820-170A-20	Sequence 20, Appl
c 35	51.2	10.9	1518	3	US-09-055-699-20	Sequence 20, Appl
c 36	51.2	10.9	1518	4	US-09-273-565-20	Sequence 20, Appl
c 37	51.2	10.9	1518	4	US-09-565-538-20	Sequence 20, Appl
c 38	51.2	10.9	2636	2	US-08-820-170A-21	Sequence 21, Appl
c 39	51.2	10.9	2636	3	US-09-055-699-21	Sequence 21, Appl
c 40	51.2	10.9	2636	4	US-09-273-565-21	Sequence 21, Appl
c 41	51.2	10.9	2636	4	US-09-565-538-21	Sequence 21, Appl
c 42	50.8	10.8	543	6	5273901-6	Patent No. 5273901
c 43	50	10.6	16442	3	US-08-781-891-208	Sequence 208, App
c 44	49.6	10.5	2369	4	US-08-910-925-2	Sequence 2, Appl
c 45	49.6	10.5	7218	1	US-08-232-463-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1  
US-09-247-155-41  
; Sequence 41, Application US/09247155A  
; Patent No. 6312922  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste  
; APPLICANT: Duclert, Aymeric  
; APPLICANT: Bouqueleret, Lydie  
; TITLE OF INVENTION: Complementary DNAs  
; FILE REFERENCE: GENSET.021A  
; CURRENT APPLICATION NUMBER: US/09/247.155A  
; CURRENT FILING DATE: 1999-02-09  
; EARLIER APPLICATION NUMBER: 60/074,121  
; EARLIER FILING DATE: 1998-02-09  
; EARLIER APPLICATION NUMBER: 60/081,563  
; EARLIER FILING DATE: 1998-04-13  
; EARLIER APPLICATION NUMBER: 60/096,116  
; EARLIER FILING DATE: 1998-08-10  
; EARLIER APPLICATION NUMBER: 60/099,273  
; EARLIER FILING DATE: 1998-10-04  
; NUMBER OF SEQ ID NOS: 182  
; SOFTWARE: Patent.pm  
; SEQ ID NO 41  
; LENGTH: 497  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 68..337  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: 68..124  
; OTHER INFORMATION: Von Heijne matrix  
; OTHER INFORMATION: score 10  
; OTHER INFORMATION: seq LVLGVSIPLVSA/QN  
; FEATURE:  
; NAME/KEY: polyA\_signal  
; LOCATION: 462..467  
; FEATURE:  
; NAME/KEY: polyA\_site  
; LOCATION: 482..497  
US-09-247-155-41

Query Match 95.2%; Score 448.4; DB 4; Length 497;  
Best Local Similarity 98.5%; Pred. No. 8.6e-126;  
Matches 463; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Oy 1 ctttgaagcatttttctgtgctccctgctatcttcagggtcaccacatgaagttcttagc 60  
|||||

```
Db 22 ctttgaagcatttttctgtctccctgatcttcagggtcaccaccatgaagtcttagc 81
Qy 61 agtctggtaactcttgggaagtttccatcttctcgtctctgcccgaatccgcacaacgc 120
Db 82 agtctcgttactcttgggaagtttccatcttctcgtctctgcccgaatccgcacaacgc 141
Qy 121 tctccagctgacagctatccagctactggtcctcgtctgatgatgaagccctgatgctga 180
Db 142 tgcctcagctgacagctatccagctactggtcctcgtctgatgatgaagccctgatgctga 201
Qy 181 aaccactgctgcaacaactcgaccactgctcctaccactgcaaacaccgcgtgc 240
Db 202 aaccactgctgcaacaactcgaccactgctcctaccactgcaaacaccgcgtgc 261
Qy 241 tctaccactgctcgaagaacattccagtttaccctaaatgggttgggagatctcccgaa 300
Db 262 tctaccactgctcgaagaacattccagtttaccctaaatgggttgggagatctcccgaa 321
Qy 301 tggtagagtgctccctgagatggaatcagcttgagcttcttgcgaattgggtcacaacta 360
Db 322 tggtagagtgctccctgagatggaatcagcttgagcttcttgcgaattgggtcacaacta 380
Qy 361 tcatgcttccctgatttccatccactacttacccttgcttacgatatcccttttatctc 420
Db 381 tcatgcttccctgatttccatccactacttacccttgcttacgatatcccttttatctc 440
Qy 421 taatcagtttattttcttccaaataaaaaataaactatgagcgagctaaaca 470
Db 441 taatcagtttattttcttccaaataaaaaataaactatgagcgagctaaaca 490

RESULT 2
US-09-253-691-3/c
; Sequence 3, Application US/09253691
; Patent No. 6124100
; GENERAL INFORMATION:
; APPLICANT: Dong Kyu JIN
; TITLE OF INVENTION: Diagnostic Method and Kit for Neuropsychiatric Diseases
; FILE REFERENCE: 1942/36
; CURRENT APPLICATION NUMBER: US/09/253,691
; EARLIER FILING DATE: 1999-02-22
; EARLIER APPLICATION NUMBER: KR 98-6, 278
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: WordPerfect 6.1/Windows
; SEQ ID NO 3
; LENGTH: 397
; TYPE: DNA
; ORGANISM: human
US-09-253-691-3

Query Match 14.6%; Score 68.6; DB 3; Length 397;
Best Local Similarity 61.5%; Pred. No. 1.9e-11;
Matches 110; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy 75 tgggaagtttccatcttctggtctctgcccagaatccgacaacagctgctccagctga 134
Db 397 TGGCCTTTCACATGATGTGAACCTGTGCTGATAGGTCCCGCTGCTGCTGCTGCTG 338
Qy 135 cgtatccagctactggtcctcgtctgatgatgaagccctgatgctgaacaccactgctg 194
Db 337 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 278
Qy 195 caaacactgcaaccactgctcctaccactgcaaacaccgcgtgcttaccactgct 253
Db 277 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 219

RESULT 3
US-09-135-994-1/c
; Sequence 1, Application US/09135994A
```

```
; Patent No. 6280938
; GENERAL INFORMATION:
; APPLICANT: Ranum et al.
; TITLE OF INVENTION: SCA7 GENE AND METHODS OF USE
; FILE REFERENCE: University of Minnesota
; CURRENT APPLICATION NUMBER: US/09/135,994A
; CURRENT FILING DATE: 1998-08-18
; EARLIER APPLICATION NUMBER: 60/056,170
; EARLIER FILING DATE: 1997-08-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 477
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-135-994-1

Query Match 13.9%; Score 65.4; DB 4; Length 477;
Best Local Similarity 63.9%; Pred. No. 1.9e-10;
Matches 99; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Qy 99 ctgcccagaatccgacaacagctgctccagctgacacgtatccagctactggtcctgctg 158
Db 300 CTGGGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 241
Qy 159 atgatgaagccctgatgctgaacaccactgctgctgcaacaactgcgaccactgctgctc 218
Db 240 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 181
Qy 219 ctaccactgcaaccaccgcgtgcttctaccactgct 253
Db 180 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 146

RESULT 4
US-08-469-802B-3/c
; Sequence 3, Application US/08469802B
; Patent No. 5741645
; GENERAL INFORMATION:
; APPLICANT: Orr, Harry T.
; APPLICANT: Ranum, Laura P.W.
; APPLICANT: Chung, Ming-yi
; APPLICANT: Zoghbi, Huda Y.
; TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
; Patent No. 5741645
; TITLE OF INVENTION: Type 1 and Method for Diagnosis
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mueeting, Raasch, Gebhardt & Schwappach, P.A.
; STREET: 119 No. 5741645th Fourth Street, Suite 203
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,802B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mueeting, Ann M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 110,00030101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-305-1217
; TELEFAX: 612-305-1225
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
```



```
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 196..1902
US-08-320-559-29

Query Match      13.8%; Score 64.8; DB 1: Length 3376;
Best Local Similarity 57.4%; Pred. No. 8.5e-10;
Matches 117; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 43 caccatgaagttcttagcagctctgtagctcttggaggtttccattcttctgtctctgc 102
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 844 CACTTTATGTTCTCTGGAGCTTTTGAAGGTTTTCCTTGCTGCTCTCATTAATTTGT 785
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 103 ccagaatccgacacagctgctccagctgacacgtatccagctactggtccctgctgatga 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 784 GAGGCTTTGAAAACTGGTACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGC 725
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 163 tgaagccctgtagctgaaacactgctgctgcaacacactgacacactgctgctctctac 222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 724 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 665
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 223 cactgcaaccaccgctgctcttac 246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 664 TACTGCTGCTGCTGCTGCTGCTGC 641

RESULT 7
US-08-545-860D-29/c
; Sequence 29, Application US/08545860D
; Patent No. 6040140
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo
; APPLICANT: Canaan, Eli
; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
; TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
; TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1 Region
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESS: No. 6040140ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545.860D
; FILING DATE: 07-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04496
; FILING DATE: 22-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10930
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/327,392
; FILING DATE: 19-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/320,559
; FILING DATE: 11-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/062,443
```

```
; FILING DATE: 14-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/971,094
; FILING DATE: 30-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,839
; FILING DATE: 27-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/805,093
; FILING DATE: 11-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca Esq., Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1262
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3376 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 196..1902
US-08-545-860D-29

Query Match      13.8%; Score 64.8; DB 3: Length 3376;
Best Local Similarity 57.4%; Pred. No. 8.5e-10;
Matches 117; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 43 caccatgaagttcttagcagctctgtagctcttggaggtttccattcttctgtctctgc 102
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 844 CACTTTATGTTCTCTGGAGCTTTTGAAGGTTTTCCTTGCTGCTCTCATTAATTTGT 785
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 103 ccagaatccgacacagctgctccagctgacacgtatccagctactggtccctgctgatga 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 784 GAGGCTTTGAAAACTGGTACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGC 725
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 163 tgaagccctgtagctgaaacactgctgctgcaacacactgacacactgctgctctctac 222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 724 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 665
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 223 cactgcaaccaccgctgctcttac 246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 664 TACTGCTGCTGCTGCTGCTGCTGC 641

RESULT 8
PCT-US94-04496-29/c
; Sequence 29, Application PC/TUS9404496
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo
; APPLICANT: Canaan, Eli
; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
; TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
; TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESS: Norris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```









```

: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 612-305-1217
: TELEFAX: 612-305-1228
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 195 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
US-08-267-803B-2

```

```

Query Match      13.5%; Score 63.4; DB 2; Length 195;
Best Local Similarity 63.4%; Pred No. 4, 8e-10;
Matches 97; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Qy 101 gccagaaatccgacacagctgctccagctgacacglatccagctactggtcctgctgat 160
Db 190 gccccggagccctgctgaggtgctgctgctgctgctgctgctgctgctgctgct 131
Qy 161 gatgaagccctgctgctgaaacacactgctgctgcaacacactgctgctgctgct 220
Db 130 gctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 71
Qy 221 accactgcaacacacgctgcttctaccactgct 253
Db 70 gctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 38

```

Search completed: October 5, 2002, 05:08:26  
Job time: 4146 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 04:00:45 ; Search time 1900.29 Seconds  
(without alignments)  
5186.787 Million cell updates/sec

Title: us-09-525-361a-23  
Perfect score: 471  
Sequence: 1 ctttgaagcattttgtctg.....aactatgagcgagtaacat 471

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_em:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vi:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htgo\_inv:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
------------	-------------	-------	--------	----	-------------

1	448.4	95.2	497	6	AR177347	AR177347 Sequence
2	448.4	95.2	578	6	AX282981	AX282981 Sequence
3	447	94.9	482	6	BD007431	BD007431 Reagent a
4	446.8	94.9	553	6	BD007432	BD007432 Reagent a
5	446.2	94.7	472	6	AX240679	AX240679 Sequence
6	446.2	94.7	473	6	AX240680	AX240680 Sequence
7	445.4	94.6	472	6	AX328498	AX328498 Sequence
8	442.2	93.9	488	6	AX067322	AX067322 Sequence
9	383.4	81.4	396	9	AF141087	AF141087 Homo sapi
10	300.8	63.9	308	6	AX240676	AX240676 Sequence
11	300.8	63.9	308	6	BD007429	BD007429 Reagent a
12	286.8	60.9	292	6	AX240677	AX240677 Sequence
13	269.8	57.3	273	6	AX328499	AX328499 Sequence
14	255.6	54.3	264	6	AX079476	AX079476 Sequence
15	255.6	54.3	266	6	AX079475	AX079475 Sequence
16	247.2	52.5	291	6	AX079474	AX079474 Sequence
17	202.8	43.1	229	6	BD007428	BD007428 Reagent a
18	187	39.7	133000	9	AC079842	AC079842 Homo sapi
19	187	39.7	188873	2	AC019223	AC019223 Homo sapi
20	187	39.7	222877	2	AC044091	AC044091 Homo sapi
21	186.4	39.6	201	6	AX240675	AX240675 Sequence
22	181.4	38.5	197	6	AX240678	AX240678 Sequence
23	181.4	38.5	197	6	BD007430	BD007430 Reagent a
24	80.8	17.2	469	10	RATSPOT1	M33976 Rat salivar
25	70.2	14.9	397	6	BD010479	BD010479 Diagnosti
26	69.6	14.8	1591	10	RNO6295	AJ006295 Rattus no
27	69.6	14.8	68409	2	AC105975	AC105975 Mus muscu
28	69.6	14.8	185363	2	AC094466	AC094466 Rattus no
29	69.4	14.7	5727	5	AY045719	AY045719 Larus arg
30	68.8	14.6	666	8	AF13050S2	AF13051 Zee may
31	68.6	14.6	46904	2	AC100911	AC100911 Mus muscu
32	67.8	14.4	98221	2	AC097839	AC097839 Rattus no
33	67.6	14.4	155780	9	HS336012	AL513498 Homo sapi
34	67.6	14.4	175695	2	AL606844	AL606844 Homo sapi
35	67.2	14.3	180668	2	AC020857	AC020857 Mus muscu
36	66.8	14.2	110000	2	AC092202_1	Continuation (2 of
37	66.8	14.2	173851	2	AC096966	AC096966 Rattus no
38	66.8	14.2	204152	2	AC092203	AC092203 Mus muscu
39	66.6	14.1	2960	3	AY057052	AY057052 Mamestra
40	66.4	14.1	624	3	AF139019	AF139019 Cepaea ne
41	66.4	14.1	236962	2	AL606742	AL606742 Mus muscu
42	66.2	14.1	22398	5	FRU271723	AJ271723 Fugu rubr
43	66.2	14.1	229380	2	AC079636	AC079636 Mus muscu
44	66	14.0	5719	10	MMCATS1	AF051726 Mus muscu
45	66	14.0	27613	9	AC002053	AC002053 Cosmid cl

ALIGNMENTS

RESULT 1	AR177347	AR177347	497 bp	DNA	linear	PAT 17-DEC-2001
LOCUS	Sequence	41 from patent	US 6312922.			
DEFINITION	AR177347					
ACCESSION	AR177347					
VERSION	AR177347.1	GI:17919702				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 497)					
AUTHORS	Edwards,J.-B.Dumas.Milne., Duciert,A. and Bougueleret,L.					
TITLE	Complementary DNAs					
JOURNAL	Patent: US 6312922-A 41 06-NOV-2001;					
FEATURES	Location/Qualifiers					
source	1..497					
BASE COUNT	122 a	138 c	90 g	147 t		
ORIGIN	/organism="unknown"					

Query Match 95.2%; Score 448.4; DB 6; Length 497;  
Best Local Similarity 98.5%; Pred. No. 1.6e-96;

Matches	463:	Conservative	0:	Mismatches	6:	Indels	1:	Gaps	1:																		
Qy	1	cttgaagcattttgtctgtccctcgatcttcaggctaccaccatgaagttcttagc	60																								
Db	22	CTTTTGAAGCATTTTGTGCTGTGCTCCCTGTGATCTTTCAAGTCAACCACTGAAGTCTTTAGC	81																								
Qy	61	agtcctgtagtactctggaggttttcacatcttcttggtctctgcccagaaatccgacaacagc	120																								
Db	82	AGTCCTGGTACTCTTGGAGGTTTCCATCTTTCTGGTCTCTGCCAGAAATCGCAACACAGC	141																								
Qy	121	tgtccagctgacacgtatccagctactcgtgctccttgctgtagatgatgaagcccctgatgctga	180																								
Db	142	TGCTCCAGCTGACACGATATCCAGCTACTTGGTCTCTGCTGATGATGAAGCCCTGATGCTGA	201																								
Qy	181	aaccactgctgctgacacaacatgcgacacactgctgctcctaccactgcaaacacacgctgc	240																								
Db	202	AACCACTGCTGCTGCAACCACTGGGACCACTGCTGCTCTCTACCACTGCAACCAACCGCTGC	261																								
Qy	241	tctaccactgctcgtgaagacatctccagttttaccacaaatgggttgaggatctcccgaa	300																								
Db	262	TTCTTACCACCTGCTCGTAAAGACATTCCAGTTTACCCTAAATGGGTGGGGATCTCCCGAA	321																								
Qy	301	lygttagagtggtccctcgagatgggaatcagcttgagcttctgcattgggtcacaacta	360																								
Db	322	TGGTAGAGTGTGCTCCCTGAGATGGAATCAGCTTGAGTCTTCTGCAATT-TGGTCAACAATA	380																								
Qy	361	tctatgcttcctgtagtttctccaaactaacttacccttgctcctacagatatccctttatctc	420																								
Db	381	TTCATGCTTCTGCTGTGATTTTCATCAACACTTACCTTGCCTTACGATATCCCTTTTATCTC	440																								
Qy	421	taatcagtttatcttcttcaataaaaaataaactatgagcgagctaaca	470																								
Db	441	TAATCAGTTTATTTCTTTTCAATTAATAAATAAATACTATACGCAAAAAA	490																								
RESULT 2																											
LOCUS	AX282981	Sequence 30 from Patent WO0175171.	578 bp	DNA	linear	PAT 03-NOV-2001																					
DEFINITION	AX282981	Sequence 30 from Patent WO0175171.																									
ACCESSION	AX282981	AX282981																									
VERSION	AX282981.1	GI:16609909																									
KEYWORDS	human.																										
SOURCE	human.																										
ORGANISM	Homo sapiens																										
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;																										
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.																										
	Houghton, R.L., Dillion, D.C., Moles, D.A., Xu, J., Zehentner, B. and																										
	Persing, D.H.																										
TITLE	Methods, compositions and kits for the detection and monitoring of																										
JOURNAL	Breast cancer																										
	Patent: WO 0175171-A 30 11-OCT-2001;																										
FEATURES	CORIXA CORPORATION (US)																										
source	Location/Qualifiers																										
	1..578																										
	/organism="Homo sapiens"																										
	/db_xref="taxon:9606"																										
BASE COUNT	206 a 137 c 88 g 147 t																										
ORIGIN																											
Query Match																											
Best Local Similarity																											
Matches 463: Conservative 0: Mismatches 6: Indels 1: Gaps 1:																											
Qy	1	cttgaagcattttgtctgtccctcgatcttcaggctaccaccatgaagttcttagc	60																								
Db	17	CTTTTGAAGCATTTTGTGCTGTGCTCCCT																									

Qy	121	tgctccagctgacacgtatccacagctactggtctctgctgatgatgaagccctcgatcgtga	180
Dy <td>137</td> <td>tgctccagctgacacgtatccacagctactggtctctgctgatgatgaagccctcgatcgtga</td> <td>196</td>	137	tgctccagctgacacgtatccacagctactggtctctgctgatgatgaagccctcgatcgtga	196
Qy <td>181</td> <td>aaccactgctctcaacaactgcgacactgctgctctcaccactgcgaaccacgcctgc</td> <td>240</td>	181	aaccactgctctcaacaactgcgacactgctgctctcaccactgcgaaccacgcctgc	240
Dy <td>197</td> <td>aaccactgctctcaacaactgcgacactgctgctctcaccactgcgaaccacgcctgc</td> <td>256</td>	197	aaccactgctctcaacaactgcgacactgctgctctcaccactgcgaaccacgcctgc	256
Qy <td>241</td> <td>ttctaccactgctcgttaagacattccagttttaccccaaatgggtgggagatctccccga</td> <td>300</td>	241	ttctaccactgctcgttaagacattccagttttaccccaaatgggtgggagatctccccga	300
Dy <td>257</td> <td>ttctaccactgctcgttaagacattccagttttaccccaaatgggtgggagatctccccga</td> <td>316</td>	257	ttctaccactgctcgttaagacattccagttttaccccaaatgggtgggagatctccccga	316
Qy <td>301</td> <td>tggtagatggtgccctgagatggaatcagcttgagcttctgcaatcggtgcgaacta</td> <td>360</td>	301	tggtagatggtgccctgagatggaatcagcttgagcttctgcaatcggtgcgaacta	360
Dy <td>317</td> <td>tggtagatggtgccctgagatggaatcagcttgagcttctgcaatcggtgcgaacta</td> <td>375</td>	317	tggtagatggtgccctgagatggaatcagcttgagcttctgcaatcggtgcgaacta	375
Qy <td>361</td> <td>ttcatgctctgctgatttccaaactacttacccttgcctacgatcctccctttatctc</td> <td>420</td>	361	ttcatgctctgctgatttccaaactacttacccttgcctacgatcctccctttatctc	420
Dy <td>376</td> <td>ttcatgctctgctgatttccaaactacttacccttgcctacgatcctccctttatctc</td> <td>435</td>	376	ttcatgctctgctgatttccaaactacttacccttgcctacgatcctccctttatctc	435
Qy <td>421</td> <td>taatcagttattttcttccaaataaaaaataactatgagcagcgtaca</td> <td>470</td>	421	taatcagttattttcttccaaataaaaaataactatgagcagcgtaca	470
Dy <td>436</td> <td>taatcagttattttcttccaaataaaaaataactatgagcagcgtaca</td> <td>485</td>	436	taatcagttattttcttccaaataaaaaataactatgagcagcgtaca	485
RESULT	3		
BD007431	BD007431	482 bp	DNA linear PAT 31-JAN-2002
LOCUS	Reagent and method useful in detecting mammary diseases.		
DEFINITION	BD007431		
ACCESSION	BD007431.1	GI:18635802	
VERSION	JP 2001503980-A/4.		
KEYWORDS	unidentified.		
SOURCE	unidentified.		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 482)		
AUTHORS	Medel,P.A.B., Cohen,M., Colpitts,T.L., Friedman,P.N., Gordon,J.,, Rapp,L.R., Granados,E.N., Hodges,S.C., Klass,M.R., Kratochvil,J.D., Rapp,L.R., Russell,J.C. and Stroupe,S.D.		
TITLE	Reagent and method useful in detecting mammary diseases		
JOURNAL	Patent: JP 2001503980-A 4 27-MAR-2001;		
COMMENT	ABBOTT LABORATORIES		
	OS Unidentified		
	PN JP 2001503980-A/4		
	PD 27-MAR-2001		
	PF 31-OCT-1997 JP 1998520795		
	PR 31-OCT-1996 US 08/742067		
	PI PATRICIA A BILLING MEDEL, MAURICE COHEN, TRACY L COLPITTS, PI POLA N FRIEDMAN,		
	PI JULIAN GORDON, EDWARD N GRANADOS, STEVEN C HODGES, PI MICHAEL R KLASS,		
	PI JOHN D KRATOCHVIL, LISA ROBERTS RAPP, JOHN C RUSSELL, PI STEVEN D STROUPE		
	PC C12N15/09,A61K38/00,A61K39/395,A61K39/395,A61P35/00,C07K14/47,		
	PC C07K16/30		
	PC C12N5/10,C12P21/02,C12Q1/68,G01N33/53,C12N15/00,C12N5/00, PC A61K37/02		
	CC Strandedness: Single;		
	CC Topology: Linear;		
	FH Key	Location/Qualifiers	
	FT source	1..482	
	FT	/organism='Unidentified'.	
FEATURES	Location/Qualifiers		
source	1..482		
	/organism='unidentified'		
	/db_xref='taxon:32644'		
BASE COUNT	109 a 136 c 90 g 145 t		
ORIGIN			
Query Match	94.98;	Score 447;	DB 6; Length 482;
Best Local Similarity	98.99;	Pred. No. 3.5e-96;	
Matches 458:	Conservative	2: Mismatches	2: Indels 1: Gaps 1:

Query Match	94.9%	Score	447	DB	6	Length	482
Best Local Similarity	98.9%	Pred. NO.	3.5e-96				
Matches	458	Conservative	2	Mismatches	2	Indels	1
						Gaps	1

```

QY 1 ctttgaagcatttttctgtgtctgtccctgatctcttcagggtaccacacatgaagtcttttagc 60
DB 18 CTTTGAAGCATTCTTGTCTGTCTGCTGATCTTCAKGTCAACACCATGAAGTTCTTAGC 77
QY 61 agtcttggtactcttggaagtttccatctttcttggtctctgtcccaagaatccgaacagc 120
DB 78 ACTCTTGCTACTCTTGGGAGTTTCCATCTTTCTGGTCTCTGGCCAGAAATCCGACAAACAGC 137
QY 121 tgcctcagctgacacgtatccagctactggtctgtctgtgatgtagaagcccttgatgctga 180
DB 138 TGGTCCAGCTGCACACGTATCCAGCTACTGGTCTGTGTGATGATGAAGCCCTTGATGCTGA 197
QY 181 aaccactgctgtgcacaaactgcgaccactgctgtctctaccactgcaacacccgctgc 240
DB 198 AACCACTGGTCTGCAACCACTGCGACCACTGCTGTCTTACCACTGCAACCAACCGCTGC 257
QY 241 tctcacactgctcgtaaagacattccagttttaccacaaatgggttgggagatctcccgaa 300
DB 258 TTTCTACCACTGCTCGTAAAGACATTTCCAGTTTACCACAAATGGGTGGGGATCTTCCGAA 317
QY 301 tggtagagtgtccctcgagatggaaatcagcttgagctctctgcaattgggtcacaacta 360
DB 318 TGGTAGAGTGTGCTCGTGAATGGAATCAGCTTGAGTCTTCTGCAATT-GGTCAACAATA 376
QY 361 ttcagtctcctgtgatttcacaaactactacotgtgctacgatataccccctttatctc 420
DB 377 TTTCACTCTCTCTGATTTTTCATCCCACTACTTACCTTGCTTACGATATCCCTTTATCTC 436
QY 421 taatcagttattttcttccaaataaaaaataaactatgagcga 463
DB 437 TAATCAGTTTATTTCTTTCAATAAAAAATAACTATGAGCAA 479

RESULT 4
HD007432 553 bp DNA linear PAT 31-JAN-2002
LOCUS
DEFINITION Reagent and method useful in detecting mammary diseases.
ACCESSION BD007432
VERSION BD007432.1 GI:18635803
KEYWORDS JP 2001503980-A/5.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 553)
AUTHORS Medel,P.A.B., Cohen,M., Colpitts,T.L., Friedman,P.N., Gordon,J.,
Granados,E.N., Hodges,S.C., Klass,M.R., Kratochvil,J.D., Rapp,L.R.,
Russell,J.C. and Stroupe,S.D.
TITLE Reagent and method useful in detecting mammary diseases
JOURNAL Patent: JP 2001503980-A 5 27-MAR-2001;
COMMENT ABBOTT LABORATORIES
OS Unidentified
PN JP 2001503980-A/5
PD 27-MAR-2001
PF 31-OCT-1997 JP 1998520795
PR 31-OCT-1996 US 08/742067
PI PATRICIA A BILLING MEDEL,MAURICE COHEN,TRACY L COLPITTS, PI
POLA N FRIEDMAN,
PI JULIAN GORDON, EDWARD N GRANADOS, STEVEN C HODGES, PI MICHAEL
R KLASS,
PI JOHN D KRATOCHVIL, LISA ROBERTS RAPP, JOHN C RUSSELL, PI
STEVEN D STROUPE
PC C12N15/09,A61K38/00,A61K39/395,A61K39/395,A61P35/00,C07K14/47,
PC C07K16/30,
PC C12N5/10,C12P21/02,C12Q1/68,G01N33/53,C12N15/00,C12N5/00, PC
A61K37/02
CC Strandedness: Single;
CC Topology: Linear;
FH Key 1..553 Location/Qualifiers
FT source /organism='Unidentified'.
FEATURES Location/Qualifiers
source 1..553

```

```

/organism='unidentified'
/db_xref='taxon:32644'
BASE COUNT 159 a 144 c 101 g 148 t 1 others
ORIGIN
Query Match 94.9%; Score 446.8; DB 6; Length 553;
Best Local Similarity 98.3%; Pred. No. 3.9e-96;
Matches 462; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
QY 1 ctttgaagcatttttctgtgtctgtccctgatctcttcagggtaccacacatgaagtcttttagc 60
DB 32 CTTTGAAGCATTCTTGTCTGTCTGCTGATCTTCAKGTCAACACCATGAAGTTCTTAGC 91
QY 61 agtcttggtactcttggaagtttccatctttcttggtctctgtcccaagaatccgaacagc 120
DB 92 AGTCTTGGTACTCTTTGGGAGTTTCCATCTTTCTGGTCTCTGCCAGAAATCCGACAAACAGC 151
QY 121 tgcctcagctgacacgtatccagctactggtctctgtgatgtagaagcccttgatgctga 180
DB 152 TGCTCCAGCTGACAGTATCCAGCTACTGGTCTCTGATGATGAAGCCCTTGATGCTGA 211
QY 181 aaccactgctgtgcacaaactgcgaccactgctgtctctaccactgcaacacccgctgc 240
DB 212 AACCACTGCTGCTGCAACCACTGCGACCACTGCTGTCTTACCACTGCAACCAACCGCTGC 271
QY 241 ttcaccactgctcgtaaagacattccagttttaccacaaatgggttgggagatctcccgaa 300
DB 272 TTTCTACCACTGCTCGTAAAGACATTTCCAGTTTACCACAAATGGGTGGGGATCTCCCGAA 331
QY 301 tggtagagtgtccctcgagatggaaatcagcttgagctctctgcaattgggtcacaacta 360
DB 332 TGGTAGAGTGTGCTCGTGAATGGAATCAGCTTGAGTCTTCTGCAATT-GGTCAACAATA 390
QY 361 ttcagtctcctgtgatttcacaaactactacotgtgctacgatataccccctttatctc 420
DB 391 TTTCACTCTCTCTGATTTTTCATCCCACTACTTACCTTGCTTACGATATCCCTTTATCTC 450
QY 421 taatcagttattttcttccaaataaaaaataaactatgagcgaacta 470
DB 451 TAATCAGTTTATTTCTTTCAATAAAAAATAACTATGAGCAACAAAAA 500

RESULT 5
AX240679 472 bp DNA linear PAT 26-SEP-2001
LOCUS
DEFINITION Sequence 5 from Patent WO0165262.
ACCESSION AX240679
VERSION AX240679.1 GI:15797662
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 472)
AUTHORS Billing-Medel,P.A., Cohen,M., Colpitts,T.L., Friedman,P.N.,
Gordon,J., Granados,E.N., Hodges,S.C., Klass,M.R., Kratochvil,J.D.,
Roberts-Rapp,L., Russell,J.C. and Stroupe,S.D.
TITLE Reagents and methods useful for detecting diseases of the breast
JOURNAL Patent: WO 0165262-A 5 07-SEP-2001;
COMMENT ABBOTT LABORATORIES (US)
FEATURES Location/Qualifiers
source 1..472
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 108 a 134 c 86 g 144 t
ORIGIN
Query Match 94.7%; Score 446.2; DB 6; Length 472;
Best Local Similarity 99.1%; Pred. No. 5.4e-96;
Matches 459; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```





```
Qy 361 ttcattgcttctggtgatttcacacactactactacctt 397
|||||
Db 360 TTATGCTTCTCTGTGATTTTCATCCAACTACTTACTT 396

RESULT 10
LOCUS AX240676 308 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 2 from Patent WO0165262.
ACCESSION AX240676
VERSION AX240676.1 GI:15797659
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Zinkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
B1ling-Medel,P.A., Cohen,M., Colpitts,T.L., Friedman,P.N.,
Gordon,J., Granados,E.N., Hodges,S.C., Klass,M.R., Kratochvil,J.D.,
Roberts-Rapp,L., Russell,J.C. and Stroupe,S.D.
Reagents and methods useful for detecting diseases of the breast
Patent: WO 0165262-A 2 07-SEP-2001;
ABBOTT LABORATORIES (US)
TITLE Location/Qualifiers
JOURNAL 1..308
FEATURES
source /organism="Homo sapiens"
BASE COUNT 63 a 94 c 65 g 86 t
ORIGIN
Query Match 63.9%; Score 300.8; DB 6; Length 308;
Best Local Similarity 99.3%; Pred. No. 2.2e-61;
Matches 302; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ctttgaagcatttttctgtgctccctgatcttcaggtcaccaccatgaagtctctagc 60
|||||
Db 5 CTTTGAAGCATTTTCTGTGCTCCCTGATCTTCAGGTCCACCACCATGAAGTCTTACG 64
Qy 61 agtctcggactcttggaggagttccatcttctgctctcctgctcctgcccagaatccgacaacagc 120
|||||
Db 65 AGTCTCGGTACTCTTGGGAGTTTCCATCTTCTGCTCTGCCCCAGAAATCCGACAACAGC 124
Qy 121 tgcctcagctgacacgtatccacgtactgctcgtcgtgatgaagccctgatgctga 180
|||||
Db 125 TGCTCCAGCTGACACGTATCCAGCTACTGGTCTGCTGATGATGAAGCCCTGATGCTGA 184
Qy 181 aaccactgctgcaaacactgacacgtactgctcctaccactgcaaccaccgctgc 240
|||||
Db 185 AACCACTGCTGTGCAACCACTGCGACCACTGCTGCTCTACCACTGCAACCAACCGCTGC 244
Qy 241 tctaccactgctcgtlaaagacattccagttttaccacaaatgggttgggagatctcccga 300
|||||
Db 245 TTCTACCACTGCTCGTAAAGACATTCACAGTTTACCCAAATGGGTGGGGATCTTCCGAA 304
Qy 301 tgggt 304
|||||
Db 305 TGGT 308

RESULT 11
LOCUS BD007429 308 bp DNA linear PAT 31-JAN-2002
DEFINITION Reagent and method useful in detecting mammary diseases.
ACCESSION BD007429
VERSION BD007429.1 GI:18635800
KEYWORDS JP 2001503980-A/2.
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 308)
AUTHORS Medel,P.A.B., Cohen,M., Colpitts,T.L., Friedman,P.N., Gordon,J.,
Granados,E.N., Hodges,S.C., Klass,M.R., Kratochvil,J.D., Rapp,L.R.,
Russell,J.C. and Stroupe,S.D.
Reagent and method useful in detecting mammary diseases
Patent: JP 2001503980-A 2 27-MAR-2001;
ABBOTT LABORATORIES
OS Unidentified
PN JP 2001503980-A/2
PD 27-MAR-2001
PF 31-OCT-1997 JP 1998520795
PR 31-OCT-1996 US 08/742067
PI PATRICIA A BILLING MEDEL,MAURICE COHEN,TRACY L COLPITTS, PI
POLA N FRIEDMAN,
PI JULIAN GORDON,EDWARD N GRANADOS,STEVEN C HODGES, PI MICHAEL
R KLASS,
PI JOHN D KRATOCHVIL,LISA ROBERTS RAPP,JOHN C RUSSELL, PI
STEVEN D STROUPE
PC C12N15/09,A61K38/00,A61K39/395,A61K39/395,A61P35/00.C07K14/47,
PC C07K16/30,
PC C12N5/10,C12P21/02,C12Q1/68,C01N33/53,C12N15/00,C12N5/00, PC
A61K37/02
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..308
/organism="Unidentified".
FEATURES
source
BASE COUNT 63 a 94 c 65 g 86 t
ORIGIN
Query Match 63.9%; Score 300.8; DB 6; Length 308;
Best Local Similarity 99.3%; Pred. No. 2.2e-61;
Matches 302; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ctttgaagcatttttctgtgctccctgatcttcaggtcaccaccatgaagtctctagc 60
|||||
Db 5 CTTTGAAGCATTTTCTGTGCTCCCTGATCTTCAGGTCCACCACCATGAAGTCTTACG 64
Qy 61 agtctcggactcttggaggagttccatcttctgctcctcctgctcctgcccagaatccgacaacagc 120
|||||
Db 65 AGTCTCGGTACTCTTGGGAGTTTCCATCTTCTGCTCTGCCCCAGAAATCCGACAACAGC 124
Qy 121 tgcctcagctgacacgtatccacgtactgctcgtcgtgatgaagccctgatgctga 180
|||||
Db 125 TGCTCCAGCTGACACGTATCCAGCTACTGGTCTGCTGATGATGAAGCCCTGATGCTGA 184
Qy 181 aaccactgctgcaaacactgacacgtactgctcctaccactgcaaccaccgctgc 240
|||||
Db 185 AACCACTGCTGTGCAACCACTGCGACCACTGCTGCTCTACCACTGCAACCAACCGCTGC 244
Qy 241 tctaccactgctcgtlaaagacattccagttttaccacaaatgggttgggagatctcccga 300
|||||
Db 245 TTCTACCACTGCTCGTAAAGACATTCACAGTTTACCCAAATGGGTGGGGATCTTCCGAA 304
Qy 301 tgggt 304
|||||
Db 305 TGGT 308

RESULT 12
LOCUS AX240677 292 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 3 from Patent WO0165262.
ACCESSION AX240677
VERSION AX240677.1 GI:15797660
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 292)
```



AUTHORS Billing-Medel, P.A., Cohen, M., Colpitts, T.L., Friedman, P.N.,  
Gordon, J., Grandos, E.N., Hodges, S.C., Klass, M.R., Kratochvil, J.D.,  
Roberts-Repp, L., Russell, J.C., and Stroupe, S.D.  
TITLE Reagents and methods useful for detecting diseases of the breast.  
JOURNAL Patent: WO 0165262-A 3 07-SEP-2001;  
ABBOTT LABORATORIES (US)  
FEATURES Location/Qualifiers  
Source 1..292  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="EST Clone 901429"  
BASE COUNT 59 a 92 c 59 g 80 t 2 others  
ORIGIN  
Query Match 60.9%; Score 286.8; DB 6; Length 292;  
Best Local Similarity 98.6%; Pred. No. 4.6e-58;  
Matches 288; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 8 gcattttgtctgtctccctgattcttcagctcaccaccatgaagttcttagcagtcctg 67  
|||||  
DB 1 GCATTTTGTCTGTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60  
|||||  
QY 68 gtactcttggagtttccatcttttctgtctctgtcccgaaatccgacacagctgctcca 127  
|||||  
DB 61 GTACTCTTTGGGAGTTTCCATCTTTCTGCTCTCTGCCAGAAATCCGACAAAGCTGCTCCA 120  
|||||  
QY 128 gctgacagatccagctacttggtctctgctgatgatgaagccctgatgctgaaccact 187  
|||||  
DB 121 CCGTACAGCTATCCAGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
|||||  
QY 188 gctgctgaacacagctgacacactgctctctaccactgcaaacacagctgcttacc 247  
|||||  
DB 181 GCTGCTGAACACTGCCAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
|||||  
QY 248 actgctgtaagacatccagctgtttaccacaaatgggtgggagatctccga 299  
|||||  
DB 241 ACTGCTGTAAGACATTNCAGTTTACCCAAATGGGTGGGATCTCCCA 292  
|||||  
RESULT 13  
AX328499  
LOCUS  
DEFINITION Sequence 2 from Patent WO0135811.  
ACCESSION AX328499  
VERSION AX328499.1 GI:18098425  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS Mack, D. and Gish, K.C.  
TITLE Novel methods of diagnosing and determining prognosis of breast  
cancer, compositions, and methods of screening for breast cancer  
modulators  
JOURNAL Patent: WO 0135811-A 2 25-MAY-2001;  
EOS Biotechnology, Inc. (US)  
FEATURES Location/Qualifiers  
Source 1..273  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 58 a 85 c 59 g 71 t  
ORIGIN  
Query Match 57.3%; Score 269.8; DB 6; Length 273;  
Best Local Similarity 99.3%; Pred. No. 5.1e-54;  
Matches 271; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 47 atgaagtcttagcagctcgtgtagcttggagtttccatcttctggtctgcccag 106  
|||||  
DB 1 ATGAAGTCTTAGCAGCTCCTGGTACTCTCTGGGAGTTTCCATCTTTCTGGTCTCTGCCAG 60  
|||||

QY 107 aatccgacacagctgtctccagctgacacgtatccagctactgctcctgctgatgatgaa 166  
|||||  
DB 61 AATCCGACACAGCTGTCTCCAGCTGACACGTATCCAGCTACTGCTGCTGCTGATGATCAA 120  
|||||  
QY 167 gcccctgatgctgaacacactgctgtgcaaacactgacacactgctgctcctaccact 226  
|||||  
DB 121 GCCCTGTATGCTGAACACACTGCTGCTCAACCACTGGGACCACTGCTGCTGCTGCTGCTGCT 180  
|||||  
QY 227 gcaacccacgctgcttaccacactgctgtaaaacacatccagtttaccacaaatggqll 286  
|||||  
DB 181 GCAACCACCGCTGCTTCTACCACTGCTCGTAACACATTCAGTTTACCCAAATGGGTT 240  
|||||  
QY 287 ggggatctcccgaaatggtagtggtgctccctga 319  
|||||  
DB 241 GGGGATCTCCGAATGGTAGAGTGTGCTCCCTGA 273  
|||||  
RESULT 14  
AX079476  
LOCUS  
DEFINITION Sequence 220 from Patent WO0107611.  
ACCESSION AX079476  
VERSION AX079476.1 GI:13159036  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS Baker, K.P., Goddard, A. and Wood, W.J.  
TITLE Human polypeptides and methods for the use thereof  
JOURNAL Patent: WO 0107611-A 220 01-FEB-2001;  
Genentech, Inc. (US)  
FEATURES Location/Qualifiers  
Source 1..264  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 54 a 87 c 53 g 70 t  
ORIGIN  
Query Match 54.3%; Score 255.6; DB 6; Length 264;  
Best Local Similarity 98.5%; Pred. No. 1.2e-50;  
Matches 258; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 ctltgaagcattttgtctgtgctccctgatcttcaggtcaccacacatgaagttctttagc 60  
|||||  
DB 3 CTTTGAAGCATTTTGTCTGTGCTCCCTGATCTTTCAGTCAACCCCAATGAAGTTCTTAGC 62  
|||||  
QY 61 agtctggtactcttggagtttccatcttctggtctctgcccagaaatccgacacagc 120  
|||||  
DB 63 AGTCCTGGTACTCTTGGGAGTTTCCATCTTCTGGTCTCTGCCAGAAATCCGACAAACAGC 122  
|||||  
QY 121 tgcctcagctgacacgtatccagctactggtctctgctgatgatgaagccctgatgctga 180  
|||||  
DB 123 TGTCTCCAGCTGACACGTATCCAGCTACTGCTGCTGCTGATGATGAAGCCCTGATGCTGA 182  
|||||  
QY 181 aaccactgctgctgcaaacactgcgacactgctgctcctaccactgcaaccacccgctgc 240  
|||||  
DB 183 AACCACTGCAACTGCAACCACTGGGACCACTGCTGCTCTTACCACCTGCAACCAACCCGCTGC 242  
|||||  
QY 241 ttctaccactgctcgtaaagac 262  
|||||  
DB 243 TTCTACCACCTGCTCGTAAAGAC 264  
|||||  
RESULT 15  
AX079475  
LOCUS  
DEFINITION Sequence 219 from Patent WO0107611.  
ACCESSION AX079475  
VERSION AX079475.1 GI:13159035

KEYWORDS  
SOURCE human.  
ORGANISM homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS 1 (bases 1 to 266)  
Haker, K.P., Goddard, A. and Wood, W.L.  
TITLE Human polypeptides and methods for the use thereof  
JOURNAL Patent: WO 0107611-A 219 01-FEB-2001;  
Genentech, Inc. (US)  
FEATURES  
Location/Qualifiers  
source  
1..266  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 54 a 88 c 54 g 70 t  
ORIGIN

Query Match 54.3%; Score 255.6; DB 6; Length 266;  
Best Local Similarity 98.5%; Pred. No. 1.2e-50;  
Matches 258; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 clttgaagcatlcttgctgctccctgatcttcaggtcaccaccatgaagtctttagc 60  
|||||  
Db 5 CTTTGAAGCATTTTCTCTGCTGCTCCCTGATCTTCAGTCCACCCCATGAAGTTCTTAGC 64  
|||||

Qy 61 aqtcctgggtactctlgggagcttccatcttctggtctctgcccagaatccgacaacagc 120  
|||||  
Db 65 AGTCTGCTACTCTGGGAGTTTCCATCTTTCTGGTCTCTGCCAGAAATCCGACAACAGC 124  
|||||

Qy 121 tgcctcagctgacagctatccagctactgctgctgctgatgatgaagccctgatgctga 180  
|||||  
Db 125 TGCTCCAGCTGACACGTATCCAGCTACTGCTGCTGCTGATGATGAAGCCCTGATGCTGA 184  
|||||

Qy 181 aacctgctgctlgcaacaacactgacaccactgctgctctaccactgcaaccaccgctgc 240  
|||||  
Db 185 AACCACTGCAACTGCAACCACCTGCGACCACTGCTGCTCTTACCCTGCAACCACCGCTGC 244  
|||||

Qy 241 tctaccactgctcgtaaaagac 262  
|||||  
Db 245 TTCTACCCTGCTCGTAAGAC 266

Search completed: October 5, 2002, 05:40:29  
Job time: 5984 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 04:39:04 ; Search time 211.95 Seconds  
(without alignments)  
3815.365 Million cell updates/sec

Title: US-09-525-361A-23  
Perfect score: 471  
Sequence: 1 cttgaacattttgtctg.....aactatgagcgagtaacat 471

Scoring table: IDENTITY\_NUC  
Gap 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_032802.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	469.4	99.7	471	21	AAA54120
2	469.4	99.7	471	22	AAF85500
3	448.4	95.2	480	20	AAH84202
4	448.4	95.2	480	21	AAH79431
5	448.4	95.2	489	20	AAH00704
6	448.4	95.2	489	20	AAH00644
7	448.4	95.2	497	20	AAZ40783
8	448.4	95.2	578	21	AAH79470
9	448.4	95.2	578	22	AAH67224

10	447	94.9	482	19	AAV31992	BS106 polynucleotl
11	447	94.9	482	20	AAH55581	BS106 consensus po
12	446.8	94.9	553	19	AAV31993	BS106 polynucleotl
13	446.8	94.9	553	20	AAH55582	BS106 consensus po
14	446.4	94.8	544	21	AAZ91765	Human breast speci
15	446.2	94.7	472	22	AAH18673	Human BS106 gene-s
16	446.2	94.7	472	22	AAH18674	Human BS106 gene-s
17	443	94.1	471	21	AAZ65099	Membrane-bound pro
18	443	94.1	471	22	AAH44245	Human PRO1160 (UNQ
19	443	94.1	471	22	AAH44245	Human anglogenesis
20	442.2	93.9	488	22	AAH44870	Human breast cance
21	300.8	63.9	308	19	AAV31990	BS106 polynucleotl
22	300.8	63.9	308	20	AAH55579	BS106 clone 893988
23	300.8	63.9	308	22	AAH18670	Human BS106 gene-s
24	298.8	63.4	501	23	AAH87721	DNA encoding novel
25	286.8	60.9	292	22	AAH18671	Human BS106 gene-s
26	271.4	57.6	273	21	AAH54121	Breast cancer prot
27	255.6	54.3	264	22	AAH93399	CDNA encoding SRT
28	255.6	54.3	266	22	AAH93398	CDNA encoding SRT
29	247.2	52.5	291	22	AAH93397	CDNA encoding SRT
30	202.8	43.1	229	19	AAV31989	BS106 polynucleotl
31	202.8	43.1	229	20	AAH55578	BS106 clone 166288
32	186.4	39.6	201	22	AAH18669	Human BS106 gene-s
33	181.4	38.5	197	19	AAV31991	BS106 polynucleotl
34	181.4	38.5	197	20	AAH55580	BS106 clone 120981
35	181.4	38.5	197	22	AAH18672	Human BS106 gene-s
36	94	20.0	660	23	AAH87720	DNA encoding novel
37	70.2	14.9	397	20	AAH89891	Spinocerebellar at
38	65.4	13.9	477	21	AAZ44307	Human SCA7 genomic
39	65	13.8	234	16	AAH84832	Spinocerebellar at
40	64.8	13.8	1326	23	AAH83104	DNA encoding novel
41	64.8	13.8	3376	16	AAH75166	AF-9 cDNA. Homo s
42	64.6	13.7	203	19	AAV30271	Glutamine rich reg
43	64.6	13.7	203	19	AAV17226	SCA2 gene CAG repe
44	64.6	13.7	1037	21	AAH55242	Exons E, C and A o
45	64.6	13.7	1472	21	AAH59241	Exons D, C, B and

#### ALIGNMENTS

RESULT 1  
AAA54120  
ID AAA54120 standard; DNA; 471 BP.  
XX  
AC AAA54120;  
XX  
DT 08-FEB-2001 (first entry)  
XX  
DE Breast cancer protein BCH1 coding sequence.  
XX  
KW Breast cancer; diagnosis; prognosis; detection; screening;  
KW antibody; oestrogen receptor; anti-oestrogen; immune response;  
KW lymph node; metastases; tumour; BCR3; BCQ8; BCQ5; BCH1; BCN1; BCN2;  
KW BCN5; BCQ5; BCQ8; BCX2; BCX3; BCA2; BCR2; BCJ7; BCY3; human; ds.  
XX  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 47..319  
FT FT /\*tag= a  
FT FT /product= BCH1 protein  
XX  
PN WO200055629-A2.  
PD 21-SEP-2000.  
XX  
PP 15-MAR-2000; 2000MO-US06952.  
XX  
PR 15-MAR-1999; 99US-0268865.  
PR 12-NOV-1999; 99US-0439878.  
PR 12-NOV-1999; 99US-0440370.  
PR 15-NOV-1999; 99US-0440493.

PR 16-NOV-1999; 99US-0440676.  
PR 16-NOV-1999; 99US-0440677.  
PR 29-NOV-1999; 99US-0450810.  
PR 02-DEC-1999; 99US-0453137.  
PR 08-MAR-2000; 2000US-0453137.  
XX  
XX  
XX (EOSB-) EOS BIOTECHNOLOGY INC.  
XX  
XX Mack D, Gish KC;  
XX  
XX WPI; 2000-638216/61.  
XX P-PSDH; AAB00184.  
XX  
XX  
XX Screening drug candidates for their ability to modulate breast cancer  
XX by contacting the drug to a cell expressing an expression profile gene  
XX and determining modulation of expression of the gene  
XX  
XX Disclosure; Fig 32; 258pp; English.  
XX  
XX New methods for screening drug candidates are described which  
XX comprise adding a drug candidate to a cell that expresses a protein  
XX selected from BClH1, BClA2, BClJ7, BClN1, BClN5, BClO2, BClQ5, BClR2, BClX2  
XX and BClY3 or their fragments and determining the effect of the drug  
XX on the expression of those proteins. Antibodies to breast cancer  
XX genes (specifically BClH1 or its fragment (BClH1p1 or BClH1p2)) are  
XX useful for inhibiting and treating breast cancer in individuals who  
XX are non-responsive to anti-oestrogen and positive for oestrogen  
XX receptor. Compositions comprising BClH1 or a nucleic acid encoding  
XX BClH1 are useful for eliciting an immune response in an individual.  
XX The antibodies are also useful for the diagnosis and prognosis of  
XX breast cancer and for screening compositions which modulate the  
XX breast cancer phenotype. The method allows rapid and simple  
XX detection of lymph node metastases.  
XX  
XX Sequence 471 BP; 108 A; 133 C; 88 G; 142 T; 0 other;

[illegible]

RESULT	2	
AAF85500		
ID	AAF85500 standard; cDNA: 471 BP.	
XX		
AC	AAF85500;	
XX		
DT	23-JUL-2001 (first entry)	
XX		
DE	Nucleotide sequence of a human breast cancer protein designated BCH1.	
XX		
KW	Breast cancer protein; BCH1; breast cancer; chromosome 12; 12q12;	
KW	anti-oestrogen therapy; ss.	
XX		
OS	Homo sapiens.	
XX		
Key	Location/Qualifiers	
FT	47..319	
CDS	/*tag= a	
FT	/product= "breast cancer protein BCH1"	
FT	47..103	
FT	/*tag= b	
XX		
WO200135811-A2.		
PN		
XX		
PD	25-MAY-2001.	
XX		
PF	16-NOV-2000; 2000WO-US31736.	
XX		
PR	16-NOV-1999; 99US-0440676.	
XX		
PA	(EOSB-) EOS BIOTECHNOLOGY INC.	
XX		
PI	Mack D, Gish KC;	
XX		
DR	WPI; 2001-343670/36.	
DR	P-PSDB; AAB83832.	
XX		
PT	BCH1 genes and protein useful for the diagnosis and treatment of breast	
PT	cancer -	
XX		
PS	Disclosure; Fig 1; 73pp; English.	

[illegible]



XX A novel isolated polypeptide comprising an immunogenic portion of a  
PT breast cancer protein useful in the detection and treatment of breast  
PT cancer -

XX Claim 4: Page 74; 95pp; English.

XX The present sequence was isolated from a breast tumour cDNA library. It  
CC is provided in a specification relating to compounds for immunotherapy  
CC and diagnosis of breast cancer. Breast tumour antigens and the  
CC polynucleotides that encode them may be used in the production of a  
CC pharmaceutical composition to be used in the treatment of breast cancer.  
CC Proliferated T cells and incubated antigen presenting cells are also  
CC required. The polypeptides and polynucleotides may also be used to  
CC produce a vaccine.

XX Sequence 480 BP; 121 A; 132 C; 87 G; 140 T; 0 other;

Query Match 95.2%; Score 448.4; DB 21; Length 480;

Best Local Similarity 98.5%; Pred. No. 2e-113;

Matches 463; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

OY 1 ctttgaagcattttgtctgtctgtccctgattctcaggtcaccaccatgaagttcttagc 60

Db 3 ctttgaagcattttgtctgtctgtccctgattctcaggtcaccaccatgaagttcttagc 62

OY 61 agtcttgtaactcttgaggagtttccatcttcttggtctctccagaaatccgacaacagc 120

Db 63 agtcttgtaactcttgaggagtttccatcttcttggtctctccagaaatccgacaacagc 122

OY 121 tgcctcagctgacacgtatccagctactggtctctgctgattgataagcccctgattgcta 180

Db 123 tgcctcagctgacacgtatccagctactggtctctgctgattgataagcccctgattgcta 182

OY 181 aaccactgctgtgacaactgcagacactgctgctctcactaccactgcaaccaccgctgc 240

Db 183 aaccactgctgtgacaactgcagacactgctgctctcactaccactgcaaccaccgctgc 242

OY 241 ttctaccactgctgtgaaagacattccagttttaccacaaatgggtgggattctcccgaa 300

Db 243 ttctaccactgctgtgaaagacattccagttttaccacaaatgggtgggattctcccgaa 302

OY 301 tggtagagtgctccctgagatggaatcagcttgagttctctgcaattgggtcacacacta 360

Db 303 tggtagagtgctccctgagatggaatcagcttgagttctctgcaattgggtcacacacta 361

OY 361 ttcatgctctgctgatttccaaactaacttaccctgctcctacgatatcccttttatctc 420

Db 362 ttcatgctctgctgatttccaaactaacttaccctgctcctacgatatcccttttatctc 421

OY 421 taatcagtttattttcttccaaataaaaaataactatgagcgagctaaaca 470

Db 422 taatcagtttattttcttccaaataaaaaataactatgagcgagctaaaca 471

RESULT 5

AAAX00704

XX ID AAAX00704 standard; DNA; 489 BP.

XX AC AAAX00704;

XX DT 25-MAR-1999 (first entry)

XX DE Human secreted protein gene 34 clone HMQAJ64.

XX Human; secreted protein; fusion protein; gene therapy; protein therapy;  
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;  
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;

KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

XX Homo sapiens.

XX WO9842738-A1.

XX 01-OCT-1998.

XX 19-MAR-1998; 98WO-US05311.

XX 30-MAY-1997; 97US-0050937.

XX 21-MAR-1997; 97US-0041276.

XX 21-MAR-1997; 97US-0041277.

XX 21-MAR-1997; 97US-0041281.

XX 21-MAR-1997; 97US-0042344.

XX 30-MAY-1997; 97US-0048069.

XX 30-MAY-1997; 97US-0048094.

XX 30-MAY-1997; 97US-0048095.

XX 30-MAY-1997; 97US-0048096.

XX 30-MAY-1997; 97US-0048099.

XX 30-MAY-1997; 97US-0048131.

XX 30-MAY-1997; 97US-0048135.

XX 30-MAY-1997; 97US-0048154.

XX 30-MAY-1997; 97US-0048160.

XX 30-MAY-1997; 97US-0048186.

XX 30-MAY-1997; 97US-0048187.

XX 30-MAY-1997; 97US-0048188.

XX 30-MAY-1997; 97US-0048350.

XX 30-MAY-1997; 97US-0048351.

XX 30-MAY-1997; 97US-0048352.

XX 30-MAY-1997; 97US-0048355.

XX 05-AUG-1997; 97US-0054804.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Brewer LA, Duan R, Ebner R, Ferrie AM, Florence KA;

XX Greene JM, Hu JS, Lafleur DM, Moore PA, NI J, Olsen HS;

XX Rosen CA, Ruben SM, Shi Y, Young P;

XX WPI; 1999-070066/06.

XX P-PSDB; AAW67900.

XX New isolated human genes and the secreted polypeptides they encode -

XX useful for diagnosis and treatment of e.g. cancers, neurological

XX disorders, immune diseases, inflammation or blood disorders

XX Claim 1; Page 254; 385pp; English.

XX This sequence represents a nucleic acid molecule which encodes a

XX secreted human protein. The gene number, and the clone it is derived

XX from, are detailed in the descriptor line. The gene can be used to

XX generate fusion proteins by linking to the gene to a human immunoglobulin

XX Fc portion (e.g. AAX00602) for increasing the stability of the fused

XX protein as compared to the human protein only.

XX The invention relates to 87 novel genes and their fragments (nucleic

XX acid sequences: AAX00611-X00724; amino acid sequences AAW67807-W68004)

XX which are useful for preventing, treating or ameliorating medical

XX conditions e.g. by protein or gene therapy. Also, pathological

XX conditions can be diagnosed by determining the amount of the new

XX polypeptides in a sample or by determining the presence of mutations in

XX the new polynucleotides. Specific uses are described for each of the 87

XX polynucleotides, based on which tissues they are most highly expressed in

XX (see AAX00611 for described uses).

XX Sequence 489 BP; 124 A; 134 C; 91 G; 140 T; 0 other;

XX

XX Query Match 95.2%; Score 448.4; DB 20; Length 489;

XX Best Local Similarity 98.5%; Pred. No. 2e-113;

XX Matches 463; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

OY 1 ctttgaagcattttgtctgtctgtccctgattctcaggtcaccaccatgaagttcttagc 60

|||||

12 ctttgaagcaattttgtctgtcctcgtgatcttcagggtcacaccatgaaattcttagc 71  
 GY agtcctggtaaccttgaggagtltcccatctttctgggtctctgccacagaatcccacaacagc 120  
 DB |||||  
 72 agtccctggtactcttgaggagtltcccatctttctggtctctgccacagaatcccacaacagc 131  
 GY |||||  
 121 tgcctccagctgacaacgtatccacgtactcgtcctcgtgatgatgaagccccctgatgctga 180  
 bb |||||  
 132 tgcctccagctgacacgtatccacgtactcgtcctcgtgatgatgaagccccctgatgctga 191  
 GY |||||  
 181 aaccactgctcgtcgcacaacactcgaccactgctcctaccactgcaaacacccgtgc 240  
 bb |||||  
 192 aaccactgctcgtcgcacaacactcgaccactgctcctaccactgcaaacacccgtgc 251  
 GY |||||  
 241 lctaccactgctcgtcgcacaacactlccagttttaccraaatgggttgggatctcccgaa 300  
 DB |||||  
 252 lctaccactgctcgtcgcacaacactlccagttltaccraaatgggttgggatctcccgaa 311  
 GY |||||  
 301 tggtagagtgctccctcgagatggaaatcagctlgagcttcttgcaattgggtccaacata 360  
 bb |||||  
 312 tggtagagtgctccctcgagatggaaatcagctlgagcttcttgcaatt-ggtccaacata 370  
 GY |||||  
 361 ttcagcttctcgtlgatttccatcccaactactlaccctlgcctacgatatcccccttatctc 420  
 DB |||||  
 371 ttcagcttctcgtlgatttccatcccaactactlaccctlgcctacgatatcccccttatctc 430  
 GY |||||  
 421 taatcagtttatittctttccaataaaaaataactatgagcggagctaaca 470  
 DB |||||  
 431 taatcagtttatittctttccaataaaaaataactatgagcaacaaaaa 480

RESULT 6  
 AAX00644  
 ID AAX00644 standard; DNA; 489 BP.  
 AC AAX00644;  
 DT  
 XX  
 XX  
 DE Human secreted protein gene 34 clone HMQAJ64.  
 XX  
 KW Human; secreted protein; fusion protein; gene therapy; protein therapy;  
 diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
 developmental abnormality; foetal deficiency; blood; allergy; renal; ds;  
 immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
 inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
 cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
 osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
 endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
 OS Homo sapiens.  
 XX  
 PN W09842738-A1.  
 XX  
 PD 01-OCT-1998.  
 XX  
 PF 19-MAR-1998; 98W0-US05311.  
 XX  
 XX 30-MAY-1997; 97US-0050937.  
 PR 21-MAR-1997; 97US-0041276.  
 PR 21-MAR-1997; 97US-0041277.  
 PR 21-MAR-1997; 97US-0041281.  
 PR 21-MAR-1997; 97US-0042344.  
 PR 30-MAY-1997; 97US-0048069.  
 PR 30-MAY-1997; 97US-0048094.  
 PR 30-MAY-1997; 97US-0048095.  
 PR 30-MAY-1997; 97US-0048096.  
 PR 30-MAY-1997; 97US-0048131.  
 PR 30-MAY-1997; 97US-0048135.  
 PR 30-MAY-1997; 97US-0048154.  
 PR 30-MAY-1997; 97US-0048160.  
 PR 30-MAY-1997; 97US-0048186.









```

QY 181 aaccactgctgcgaacaactgcgaccactgctgctcctaccactgcgaaccacgcgtgc 240
Db 198 aaccactgctgcgaacaactgcgaccactgctgctcctaccactgcgaaccacgcgtgc 257
QY 241 tctaccactgctgcgaagaacatccagttttacccaatgggttgaggatcccca 300
Db 258 tctaccactgctgcgaagaacatccagttttacccaatgggttgaggatcccca 317
QY 301 tggtagagtgctccctgagatggaaatcagcttgagttcttgcgaattgggtcacaaacta 360
Db 318 tggtagagtgctccctgagatggaaatcagcttgagttcttgcgaattgggtcacaaacta 376
QY 361 ttcacgtcttcctgatttcacacactacttacccttgctacgatatccctttatctc 420
Db 377 ttcacgtcttcctgatttcacacactacttacccttgctacgatatccctttatctc 436
QY 421 taatcagttatttttttccaaataaaaaataactatgagcga 463
Db 437 taatcagttatttttttccaaataaaaaataactatgagcga 479

RESULT 12
AAV31993
ID AAV31993 standard; DNA; 553 BP.
AC AAV31993;
DT 25-SEP-1998 (first entry)
DE BS106 polynucleotide clone 1652885 (ii).
KW ss: human; BS106; breast tissue gene; breast cancer; detection marker.
OS Homo sapiens.
PN WO9818945-A1.
PD 07-MAY-1998.
PF 31-OCT-1997; 97MO-US19836.
PR 31-OCT-1996; 96US-0742067.
XX (ABBO ) ABBOTT LAB.
PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN;
PI Gordon J, Granados EN, Hodges SC, Klass MR, Kratochvil JD;
PI Roberts-Rapp L, Russell JC, Stroupe SD;
DR WPI; 1998-272234/24.
PT Detecting BS106 gene products - useful for developing products for
PT detecting, staging, preventing, treating or determining
PT predisposition to e.g. breast cancer
XX Claim 1; Page 90; 114pp; English.
XX BS106 is a breast tissue gene with which breast cancer and related
CC diseases are associated. The BS106 polynucleotides AAV31989-V31997 can
CC be detected by BS106 specific polynucleotides or complements acting as
CC useful markers for detection methods. The products and methods can be
CC used for detecting, diagnosing, staging, preventing or treating or
CC determining predisposition to diseases or conditions of the breast such
CC as breast cancer.
XX Sequence 553 BP; 159 A; 144 C; 101 G; 148 T; 1 other;
SQ

Query Match 94.9%; Score 446.8; DB 19; Length 553;
Best Local Similarity 98.3%; Pred. No. 5.9e-113;
Matches 462; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
QY 1 ctttgaagcattttgtctgctcctgatcttcacacaccatgaagtcttagc 60

```

```

Db 32 ctttgaagcattttgtctgctcctgatcttcacacaccatgaagtcttagc 91
QY 61 agtctggtactcttggagtttccatcttttctggtctctgccagaatccgacaacagc 120
Db 92 agtctggtactcttggagtttccatcttttctggtctctgccagaatccgacaacagc 151
QY 121 tgcctcagctgcacagtattccagctactggtcctgctgatgatgaagcccttgatgctga 180
Db 152 tgcctcagctgcacagtattccagctactggtcctgctgatgatgaagcccttgatgctga 211
QY 181 aaccactgctgcgaacaactgcgaccactgctgctcctaccactgcgaaccacgcgtgc 240
Db 212 aaccactgctgcgaacaactgcgaccactgctgctcctaccactgcgaaccacgcgtgc 271
QY 241 tctaccactgctgcgaagaacatccagttttacccaatgggttgaggatcccca 300
Db 272 tctaccactgctgcgaagaacatccagttttacccaatgggttgaggatcccca 331
QY 301 tggtagagtgctccctgagatggaaatcagcttgagttcttgcgaattgggtcacaaacta 360
Db 332 tggtagagtgctccctgagatggaaatcagcttgagttcttgcgaattgggtcacaaacta 390
QY 361 ttcacgtcttcctgatttcacacactacttacccttgctacgatatccctttatctc 420
Db 391 ttcacgtcttcctgatttcacacactacttacccttgctacgatatccctttatctc 450
QY 421 taatcagttatttttttccaaataaaaaataactatgagcga 470
Db 451 taatcagttatttttttccaaataaaaaataactatgagcga 500

```

## RESULT 13

AAV55582

ID AAV55582 standard; DNA; 553 BP.

XX AC AAV55582;

XX DT 26-JUL-1999 (first entry)

XX DE BS106 consensus polynucleotide sequence.

XX KW BS106; breast; marker; detection; diagnosis; breast cancer; ss.

XX OS Homo sapiens.

XX PN WO9233230-A1.

XX PD 14-MAY-1999.

XX PF 19-OCT-1998; 98WO-US22020.

XX PR 31-OCT-1997; 97US-0962094.

XX PA (ABBO ) ABBOTT LAB.

XX PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN;

XX PI Gordon J, Granados EN, Hodges SC, Klass MR, Kratochvil JD;

XX PI Roberts-Rapp L, Russell JC, Stroupe SD;

XX DR WPI; 1999-326991/27.

XX DR P-PSDB; AAY13466.

XX PT New isolated BS106 nucleic acids

XX PS Claim 11; Page 89; 105pp; English.

Sequences AAV55578-82 represent overlapping polynucleotide sequences of BS106 clones isolated from the breast tissue. The BS106 nucleic acids correspond to a breast tissue gene and can be used as markers for breast tissue disease. The products can be used for detecting, diagnosing, staging, monitoring, prognosticating, preventing or treating, or determining diseases or conditions associated with BS106, especially



```
PF 28-FEB-2001: 2001WO-US06516.
XX
PR 29-FEB-2000: 2000US-0516444.
XX
XX (ABBO ) ABBOTT LAB.
XX
PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Roberts-Rapp L;
PI Russell JC, Stroupe SD;
XX
DR WPI: 2001-596773/67.
XX
XX A gene or its fragment which codes for a BSI06 polypeptide, useful for
PT the detection of a breast disease such as breast cancer.
XX
XX Example 1; Fig 1; 158pp; English.
XX
XX The invention relates to human BSI06-specific polypeptides and
CC polynucleotides. The BSI06 polypeptides and antibodies are useful
CC for detecting, diagnosing, staging, monitoring, prognosticating,
CC preventing, treating or determining the predisposition of an
CC individual to diseases and conditions of the breast such as breast
CC cancer. They are also useful in the treatment of tumours or
CC metastases. Polynucleotides of the invention are useful in drug
CC screening and gene therapy. The present sequence is human BSI06
CC gene specific full length cDNA, expressed sequence tag (EST) clone
XX 16628851nh.
XX
SQ Sequence 472 BP; 108 A; 134 C; 86 G; 144 T; 0 other;
```

Query Match 94.7%; Score 446.2; DB 22; Length 472;  
Best Local Similarity 99.1%; Pred. No. 8e-113;  
Matches 459; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```
OY 1 ctttgaagcatttttctgtgtcctcctgatcttcaggccaccacccatgaagtctcttagc 60
Db 9 ctttgaagcatttttctgtgtcctcctgatcttcaggccaccacccatgaagtctcttagc 68
OY 61 agtcctgttactcttgggagttccatcttctgtctgtcctcctgatcttcaggccaccacccacg 120
Db 69 agtcctgttactcttgggagttccatcttctgtctgtcctcctgatcttcaggccaccacccacg 128
OY 121 tgcctcagctgacagctatccagctactggtcctcctgatgatgaagccctcgtatgctga 180
Db 129 tgcctcagctgacagctatccagctactggtcctcctgatgatgaagccctcgtatgctga 188
OY 181 aaccactgctgtgcaacaactgcgaccactgctcctcctaccactgcaaccacccgctgc 240
Db 189 aaccactgctgtgcaaccactgcgaccactgctcctcctaccactgcaaccacccgctgc 248
OY 241 ttctaccactgctgtaagacatctccagttttaccctcccaaatgggttggggatctcccga 300
Db 249 ttctaccactgctgtaagacatctccagttttaccctcccaaatgggttggggatctcccga 308
OY 301 tggtagagtgtcctcctgagatgggaatcagcttgagttctcctgcaattgggtcacacaacta 360
Db 309 tggtagagtgtcctcctgagatgggaatcagcttgagttctcctgcaatt-ggtcacacaacta 367
OY 361 ttcatgcttctcctgatttccatcccaactacttaccttgcctacgatatccctttatctc 420
Db 368 ttcatgcttctcctgatttccatcccaactacttaccttgcctacgatatccctttatctc 427
OY 421 taatcagttattttcttcaataaaaaataaataaatatgagcga 463
Db 428 taatcagttattttcttcaataaaaaataaataaatatgagcga 470
```

Search completed: October 5, 2002, 05:44:20  
Job time: 3916 sec